PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 39/395

A1

(11) International Publication Number: WO 98/19704

(43) International Publication Date: 14 May 1998 (14.05.98)

(21) International Application Number: PCT/US97/19203

(22) International Filing Date: 23 October 1997 (23.10.97)

(30) Priority Data:
60/030,149 1 November 1996 (01.11.96) US

(71) Applicant (for all designated States except US): SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): DEEN, Keith, Charles [US/US]; 210 Kathleen Way, Glenmoore, PA 19343 (US). DILLON, Susan, Beth [US/US]; 1209 Tullamore Circle, Chester Springs, PA 19425 (US). PORTER, Terence, Graham [GB/US]; 7634 Levis Road, Cheltenham, PA 19012 (US). SWEET, Raymond, Whitney [-/US]; 108 Edgehill Road, Bala Cynwyd, PA 19004 (US).

(74) Agents: GEIGER, Kathleen, W. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: HUMAN MONOCLONAL ANTIBODIES

(57) Abstract

This invention relates to novel human monoclonal antibodies (mAbs) and to the genes encoding same. More specifically, this invention relates to human monoclonal antibodies specifically reactive with an epitope of the fusion (F) protein of Respiratory Syncytial Virus (RSV). Such antibodies are useful for the therapeutic and/or prophylactic treatment of RSV infection in human patients, particularly infants and young children.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN MONOCLONAL ANTIBODIES

Field of the Invention:

This invention relates to novel human monoclonal antibodies (mAbs) and to the genes encoding same. More specifically, this invention relates to human monoclonal antibodies specifically reactive with an epitope of the fusion (F) protein of Respiratory Syncytial Virus (RSV). Such antibodies are useful for the therapeutic and/or prophylactic treatment of RSV infection in human patients, particularly infants and young children.

Background of the Invention:

Respiratory syncytial virus (RSV) is the major cause of lower respiratory disease in children, giving rise to predictable annual epidemics of bronchiolitis and pneumonia in children worldwide. The virus is highly contagious, and infections can occur at any age. Immunity to RSV appears to be short-lived, thus reinfections are frequent. Zero to 2 year old infants are the most susceptible and represent the primary affected population. In this group, 1 out of 5 will develop lower respiratory (below larynx) disease upon infection and this ratio stays the same upon reinfection. Depending on age, environment and other associated factors, hospitalization is required in 1-3% of cases of RSV infection and is usually of long duration (up to 3 weeks). The high morbidity of RSV infection, especially in infancy, has also been implicated in the development of respiratory problems later in life. Mortality is generally very low in more developed countries, but much higher in less developed countries and in certain risk groups such as children with heart/lung disease, making prophylactic treatment desirable for these groups of children.

A vaccine for RSV infection is not currently available. Severe safety issues surrounding an attenuated whole virus vaccine tested in the 1960s, as well as the potential of induced immunopathology associated with the newer candidate subunit vaccines make the prospects of a vaccine in the near future appear remote. To date one drug therapy, Ribavirin, a broad spectrum antiviral, has been approved. Ribavirin has gained only minimal acceptance owing to problems of administration,

mild toxicity and questionable efficacy. In the majority of cases, hospitalized children receive no drug therapy and receive only intensive supportive care which is extremely costly. It is clear that there is a need for a safe, effective and easily administered drug for the treatment of RSV infection.

The feasibility of passive antibody treatment/protection against RSV has been well established using animal models. Most of the earlier passive transfer studies in animals against infectious agents, including RSV, utilized murine mABs. Recently, the FDA has approved for use intravenous gammaglobulins (IVIG) isolated from pooled human sera. Initial reports from this study had been encouraging (Groothuis, J. R. et al., Antimicrob. Agents Chemo. 35(7):1469-1473 (1991)). However, generic shortcomings of IVIGs exist and include, without limitation, the fact that such products are human blood derived and grams of antibody often need to be administered to achieve an effective dose.

Alternatively, monoclonal antibodies have been employed. The advantages of such an approach include: a higher concentration of specific antibody can be achieved thereby reducing the amount of globulin required to be given; the reliance on direct blood products can be eliminated; the levels of antibody in the preparation can be more uniformly controlled and the routes of administration can be extended. While passive immunotherapy employing monoclonal antibodies from a heterologous species (e.g., murine) has been suggested (See: PCT Application PCT/US94/08699, Publication No. WO 95/04081), one alternative to reduce the risk of an undesirable immune response on the part of the patient directed against the foreign antibody is to employ "humanized" antibodies. These antibodies are substantially of human origin, with only the Complementarity Determining Regions (CDRs) being of non-human origin. Particularly useful examples of this approach are disclosed in PCT Application PCT/GB91/01554, Publication No. WO 92/04381 and PCT Application PCT/GB93/00725, Publication No. WO93/20210.

A second and more preferred approach is to employ fully human mAbs. Unfortunately, there have been few successes in producing human monoclonal antibodies through classic hybridoma technology. Indeed, acceptable human fusion partners have not been identified and murine myeloma fusion partners do not work

well with human cells, yielding unstable and low producing hybridoma lines. However, recent advances in molecular biology and immunology make it now possible to isolate human mABs, particularly directed against foreign infectious agents, as explained in greater detail below.

Comprehensive details concerning RSV infection and its clinical features can be obtained from excellent recent reviews by McIntosh, K. and R. M. Chanock, In: "Respiratory Syncytial Virus", Ch. 38, B.N. Fields ed., Raven Press (1990) and Hall, C.B., In: "Textbook of Pediatric Disease" Feigin and Cherry, eds., W.B. Saunders, pgs 1247-1268 (1987). RSV, belonging to the family paramyoxoviridae, is a negative-strand unsegmented RNA virus with properties similar to those of the paramyxoviruses. It has, however been placed in a separate genus Pneumovirus, based on morphologic differences and lack of hemagglutinin and neuraminidase activities. RSV is pleomorphic and ranges in size from 150-300 nm in diameter. The virus matures by budding from the outer membrane of a cell and virions appear as membrane-bound particles with short, closely spaced projections or "spikes". The RNA genome encodes 10 unique viral polypeptides ranging in size from 9.5 kDa to 160 kDa (Huang, Y. T. and G. W.Wertz, J. Virol. 43:150-157 (1982)). Seven proteins (F, G, N, P, L, M, M2) are present in RSV virions and at least three proteins (F, G, and SH) are expressed on the surface of infected cells. The F protein has been conclusively identified as the protein responsible for cell fusion since specific antibodies to this protein inhibit syncytia formation *in vitro* and cells infected with vaccinia virus expressing recombinant F protein form syncytia in the absence of other RSV virus proteins. In contrast, antibodies to the G protein do not block syncytia formation but prevent attachment of the virus to cells.

RSV can be divided into two antigenically distinct subgroups, (A & B) (Mufson, M. A. et al., J. Gen'l. Virol. 66:2111-2124 (1985)). This antigenic dimorphism is linked primarily to the surface attachment (G) glycoprotein (Johnson, R. A. et al., Proc. Nat'l. Acad. Sci. USA 84:5625-5629 (1987)). Strains of both group A and B circulate simultaneously, but the proportion of each may vary unpredictably from year to year. An effective therapy must therefore target both subgroups of the

virus and this is the reason for the selection of the highly conserved surface fusion (F) protein as target antigen for mAb therapy as will be discussed later.

RSV is distributed worldwide. One of the most remarkable features of the epidemiology of RSV virus, as mentioned above, is the consistent pattern of infection and disease. Other respiratory viruses cause epidemics at irregular intervals or exhibit a mixed endemic/epidemic pattern, but RSV is the only respiratory viral pathogen that produces a sizable epidemic every year in large urban centers. In the temperate areas of the world, RSV epidemics have occurred primarily in the late fall, winter or spring but never during the summer. The occurrence and spread of infection within a community is characteristic and easily diagnosed, leading to sharp rises in cases of bronchiolitis and pediatric pneumonia and the number of hospital admissions of young children with acute lower respiratory tract disease. Other respiratory viral agents that occur in outbreaks are rarely present at the same time as RSV . Primary RSV infection occurs in the very young and virtually all children have been infected before they have entered school. By 1 year of age, 25-50% of infants have specific antibodies as a result of natural infection and this is close to 100% by age 4-5. Age, sex, socioeconomic and environmental factors can all influence the severity of disease. With current intensive care in the U.S., overall mortality for normal subjects is low (less than 2% of hospitalized subjects) but can be much higher in infants with underlying cardiac condition (cyanotic congenital heart disease) or respiratory disease (bronchopulmonary dysplasia) where the progression of symptoms may be rapid. For instance, mortality in infants with cyanotic congenital heart disease has been reported to be as high as 37%. In premature infants apneic spells due to RSV infection may occur and, in rare cases, cause neurologic or systemic damage. Severe lower respiratory tract illness (bronchiolitis and pneumonia) is most common in patients under six months of age. Infants who have apparently recovered completely from this illness may display symptomatic respiratory abnormalities for years (recurrent wheezing, decreased pulmonary function, recurrent cough, asthma, and bronchitis).

The mechanisms by which the immune system protects against RSV infection and reinfection are not well understood. It is clear, however, that immunity

is only partially protective since reinfection is common at all ages, and sometimes occurs in infants only weeks after recovery from a primary infection. Both serum and secretory antibodies (IgA) have been detected in response to RSV infection in adults as well as in very young infants. However, the titers of serum antibodies to the viral F or G glycoprotein, as well as of neutralizing antibodies found in infants (1-8 months of age) are 15-25% of those found in older subjects. These reduced titers may contribute to the increased incidence of serious infection in younger children.

Evidence for the role of serum antibodies in protection against RSV virus has emerged from epidemiological as well as animal studies. In adults exposed naturally to the virus, susceptibility correlated well with low serum antibody level. In infants, titers of maternally transmitted antibodies correlate with resistance to serious disease (Glezen, W.P. et al., J. Pediatr. 98:708-715 (1981)). Other studies show that the incidence and severity of lower respiratory tract involvement is diminished in the presence of high serum antibody (McIntosh, K. et al., J. Infect. Dis. 138:24-32 (1978)) and high titers of passively administered serum neutralizing antibodies have been shown to be protective in a cotton rat model, of RSV infection (Prince, G. A. et al., Virus Res. 3:193-206 (1985)).

Children lacking cell-mediated immunity are unable to cease their infection and shed virus for many months in contrast to children with normal immune systems. Similarly, nude mice infected with RSV virus persistently shed virus. These mice can be cured by adoptive transfer of primed T cells (Cannon, M. J. et al., Immunology 62:133-138 (1987)).

In summary, it appears that both cellular and humoral immunity are involved in protection against infection, reinfection and RSV disease and that although antigenic variation is limited, protective immunity is not complete even after multiple exposures.

This invention relates to the use of human mABs specific for the F protein of RSV virus to passively treat or prevent infection. The use of passive antibody therapy in humans is well documented and is being used to treat other infectious diseases such as hepatitis and cytomegalovirus. Clinical trials are also on-going to evaluate the efficacy of humanized antibodies for treatment of RSV infection in

young children. Studies in animals have clearly demonstrated that polyclonal and monoclonal antibody against both F and G glycoprotein can confer passive protection in RSV virus infection when given prophylactically or therapeutically (Prince, et al., supra). In these studies, passive transfer of neutralizing F or G mAbs to mice, cotton rats or monkeys, significantly reduce or completely prevent replication of the RSV virus in the lungs.

The induction of neutralizing antibodies to RSV virus appears to be limited to the F and G surface glycoproteins. Of these two proteins, the F protein is the major target for cross-reactive neutralizing antibodies associated with protection against different strains of RSV virus. In addition, experimental vaccination of mice or cotton rats with F protein also results in cross protection. The antigenic relatedness of the F protein across strains and subgroups of the virus is reflected in its high degree of homology at the amino acid level. In contrast, in the two subgroups and various strains of RSV, antigenic dimorphism was linked primarily to the G glycoprotein. The F protein has a predicted molecular weight of 68-70 kDa; a signal peptide at its N-terminus; a membrane anchor domain at its C terminus; and is cleaved proteolytically in the infected cell prior to virion assembly to yield disulfide linked F_2 and F_1. Five neutralizing epitopes have been identified within the F protein sequence and map to residues 205-225; 259-278; 289-299: 483-488 and 417-438. Studies to determine the frequency of sequence diversion in the F protein showed that the majority of the neutralizing epitopes were conserved in all of the 23 strains of RSV virus isolated in Australia, Europe, and regions of the U.S. over a period of thirty years. In another study, seroresponses of forty three infants and young children to primary infection with subgroup A or a subgroup B strain showed that responses to homologous and heterologous F antigens were not significantly different, while the G proteins of the subgroup A and B strains were quite unrelated. Moreover, antibody inhibition of virus-mediated cell fusion *in vitro* versus inhibition of infection correlates best with protection in animal models and fusion inhibition is primarily restricted to F protein specific antibodies. Clearly, the F protein is the more important target for antibody therapy.

Fully human mAbs to RSV F protein remain a desirable option for the treatment of this disease. Although some success has been reported in obtaining fragments of such mAbs (Barbas, C.F. et al., Proc. Nat'l. Acad. Sci. USA 89:10164-10168 (1992);Crowe, J. E. et al., Proc. Nat'l. Acad. Sci. USA 91: 1386-1390 (1994) and PCT application number PCT/US93/08786, published as WO94/06448, March 31, 1994)), the achievement of such results is not straight forward and novel human mABs as described herein, when and however obtained, are particularly useful alone or in combination with existing molecules to form immunotherapeutic compositions. This invention relates to one such group of human mAbs.

Brief Description of the Invention:

This invention relates to fully human monoclonal antibodies and functional fragments thereof specifically reactive with an F protein epitope of RSV and capable of neutralizing RSV infection.

In a related aspect, the present invention provides modifications to neutralizing Fab fragments or $F(ab')_2$ fragments specific for the F protein of RSV produced by random combinatorial cloning of human antibody sequences and isolated from a filamentous phage Fab display library.

In still another aspect, there is provided a reshaped human antibody containing human heavy and light chain constant regions from a first human donor and heavy and light chain variable regions or the CDRs thereof derived from human neutralizing monoclonal antibodies for the F protein of RSV derived from a second human donor.

In yet another aspect, the present invention provides a pharmaceutical composition which contains one (or more) altered antibodies and a pharmaceutically acceptable carrier.

In a further aspect, the present invention provides a method for passive immunotherapy of RSV disease in a human by administering to said human an effective amount of the pharmaceutical composition of the invention for the prophylatic or therapeutic treatment of RSV infection.

In yet another aspect, the present invention provides methods for, and components useful in, the recombinant production of human and altered antibodies (e.g., engineered antibodies, CDRs, Fab or $F(ab)_2$ fragments, or analogs thereof) which are derived from human neutralizing monoclonal antibodies (mAbs) for F protein of RSV. These components include isolated nucleic acid sequences encoding same, recombinant plasmids containing the nucleic acid sequences under the control of selected regulatory sequences which are capable of directing the expression thereof in host cells (preferably mammalian) transfected with the recombinant plasmids. The production method involves culturing a transfected host cell line of the present invention under conditions such that the human or altered antibody is expressed in said cells and isolating the expressed product therefrom.

In yet another aspect of the invention is a method to diagnose the presence of RSV in a human which comprises contacting a sample of biological fluid with the human antibodies and altered antibodies of the instant invention and assaying for the occurrence of binding between said human antibody (or altered antibody) and RSV.

In yet another embodiment of the invention is a pharmaceutical composition comprising at least one dose of an immunotherapeutically effective amount of the monoclonal antibody of this invention in combination with at least one additional monoclonal antibody. Especially, when the additional monoclonal antibody is an anti-RSV antibody distinguished from the subject antibody of by virtue of being reactive with a different epitope of the RSV F protein antigen.

Other aspects and advantages of the present invention are described further in the detailed description and the preferred embodiments thereof.

Brief Description of the Drawings:

Figure 1 illustrates the cloning strategy used for the construction of the Hu 19A monoclonal antibody. The heavy chain V region was cloned into the PCD derivative vector as a *Xho*I - *Bsp*120I fragment. The entire light chain V and C regions were cloned into the PCN derivative vector as a *Sac*I - *Xba*I fragment. Details are described in the hereinbelow.

Figure 2 provides a comparison of the heavy chain amino acid sequences of various monoclonal antibodies of this invention. The amino acid sequences of the heavy chains for the A, B, C and D constructs are shown (SEQ ID NOS: 5, 6, 7 and 8, respectively). Numbering of the residues is based on the germline (GL) gene Dp58 (SEQ ID No: 4), beginning at the mature processed amino terminus and ending at CDR3. The "-" indicates identity to the preceding sequence (eg., C compared to B). Sequence A has an amino acid insertion between positions 4 and 5 due to the cloning strategy utilized by Barbas et al. (Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992), PCT publication WO94/06448). Bold residues correspond to the leader region, and to CDRs 1-3. The underlined sequence in CDR2 identifies the N-linked glycosylation site in versions A and B that was mutated in version C. Residues P14 and G15, marked with an "*" were listed as L and A, respectively in the published sequence (Barbas et al., supra).

Figure 3 provides a comparison of the light chain amino acid sequences of various monoclonal antibodies of this invention. The amino acid sequences of the light chains for the A, B, C and D constructs are shown (SEQ ID NOS: 10, 11, 12 and 13). Numbering of the residues in the Vκ region is based on the germline (GL) gene Dpk9 (SEQ ID NO: 9), beginning at the mature processed amino terminus and ending at CDR3; but for reference to framework 4, the actual numbering is also shown for Hu19ALc. As in Fig. 2, the "-" indicates identity to the preceding sequence. The G at position 97 in framework 4 of Hu19A, marked with an "*", was listed as E in the published sequence (see text). Sequence A has a two amino acid deletion at residues 1 and 2 due to the cloning strategy. Bold residues correspond to the leader region, and to CDRs 1-3. The κ constant region is shown for constructs A and B in comparison to the germline gene. The L mutation near the C-terminus was corrected in version C (See; Figure 3, SEQ ID NO:13).

Figure 4 illustrates the DNA sequences of plasmids for the expression of the Hu19 mAB heavy and light chains. Figure 4A is the DNA sequence of Hu19AHcpcd (SEQ ID NO:14). The start of translation, leader peptide, amino-terminal processing site (SEQ ID NO:15), carboxy terminus of the 19A heavy chain

(SEQ ID NO: 16) and *Eco* RI restriction endonuclease cleavage site are shown. Figure 4B is the DNA sequence of Hu19ALcpcn (SEQ ID NO: 17), and shows the corresponding features for the light chain and the *Xba* I restriction site following the end of the coding region for the light chain (SEQ ID NO'S: 18, 19). Figure 4C is the DNA sequence of the coding region of the heavy chain of plasmid Hu19BHcpcd (SEQ ID NO'S 20,21). Figure 4D is the DNA sequence of the coding region for the light chain of plasmid Hu19BLcpcn (SEQ ID NO:22,23 & 24). Figure 4E is the DNA sequence of the coding region of the heavy chain of the plasmid Hu19CHcpcd (SEQ ID NO'S 25,26). Figure 4F is the DNA sequence of the coding sequence of the heavy chain of plasmid Hu19DHcpcd (SEQ ID NO:'S 27,28). Figure 4G is the DNA sequence of the coding region of the light chain of plasmid Hu19CLcpcn (SEQ ID NO'S: 29, 30). In Figures 4C-G, bolded residues indicate differences from the full vector sequences for Hu19AHcpcd and Hu19ALc shown in Figures 4A and 4B, respectively.

Figure 5 illustrates a Coomassie stained SDS-PAGE gel of Hu19B and Hu19C under reducing conditions.

Figure 6 illustrates the separation of Hu19 Glycovarients by anion exchange chromatography.

Figure 7 illustrates SDS-PAGE analysis of Hu19B Fab glycovarients.

Detailed Description of the Invention:

This invention provides useful human monoclonal antibodies (and fragments thereof) reactive with the F protein of RSV, isolated nucleic acids encoding same and various means for their recombinant production as well as therapeutic, prophylactic and diagnostic uses of such antibodies and fragments thereof.

I. Definitions.

As used in this specification and the claims, the following terms are defined as follows:

"Altered antibody" refers to a protein encoded by an altered immunoglobulin coding region, which may be obtained by expression in a selected host cell. Such

altered antibodies are engineered antibodies (e.g., chimeric, humanized, or reshaped or immunologically edited human antibodies) or fragments thereof lacking all or part of an immunoglobulin constant region, e.g., Fv, Fab, or $F(ab')_2$ and the like.

"Altered immunoglobulin coding region" refers to a nucleic acid sequence encoding an altered antibody of the invention or a fragment thereof.

"Reshaped human antibody" refers to an altered antibody in which minimally at least one CDR from a first human monoclonal donor antibody is substituted for a CDR in a second human acceptor antibody. Preferrably all six CDRs are replaced. More preferrably an entire antigen combining region (e.g., Fv, Fab or $F(ab')_2$) from a first human donor monoclonal antibody is substituted for the corresponding region in a second human acceptor monoclonal antibody. Most preferrably the Fab region from a first human donor is operatively linked to the appropriate constant regions of a second human acceptor antibody to form a full length monoclonal antibody. The reshaped human monoclonal antibodies designated herein as Hu19A, Hu19B, Hu19C and Hu19D are defined as reshaped human antibodies comprising a light chain amino acid sequence selected from Sequences 19A, 19B, 19C and 19D of Figure 3 and a heavy chain amino acid sequence selected from Sequences 19A, 19B, 19C and 19D of Figure 2, or functional partial sequences thereof.

"First immunoglobulin partner" refers to a nucleic acid sequence encoding a human framework or human immunoglobulin variable region in which the native (or naturally-occurring) CDR-encoding regions are replaced by the CDR-encoding regions of a donor human antibody. The human variable region can be an immunoglobulin heavy chain, a light chain (or both chains), an analog or functional fragments thereof. Such CDR regions, located within the variable region of antibodies (immunoglobulins) can be determined by known methods in the art. For example Kabat et al. (Sequences of Proteins of Immunological Interest, 4th Ed., U.S. Department of Health and Human Services, National Institutes of Health (1987)) disclose rules for locating CDRs. In addition, computer programs are known which are useful for identifying CDR regions/structures.

"Second fusion partner" refers to another nucleotide sequence encoding a protein or peptide to which the first immunoglobulin partner is fused in frame or by

means of an optional conventional linker sequence (i.e., operatively linked). Preferably the fusion partner is an immunoglobulin gene and when so, it is referred to as a "second immunoglobulin partner". The second immunoglobulin partner may include a nucleic acid sequence encoding the entire constant region for the same (i.e., homologous - the first and second altered antibodies are derived from the same source) or an additional (i.e., heterologous) antibody of interest. It may be an immunoglobulin heavy chain or light chain (or both chains as part of a single polypeptide). The second immunoglobulin partner is not limited to a particular immunoglobulin class or isotype. In addition, the second immunoglobulin partner may comprise part of an immunoglobulin constant region, such as found in a Fab, or $F(ab)_2$ (i.e., a discrete part of an appropriate human constant region or framework region). A second fusion partner may also comprise a sequence encoding an integral membrane protein exposed on the outer surface of a host cell, e.g., as part of a phage display library, or a sequence encoding a protein for analytical or diagnostic detection, e.g., horseradish peroxidase, β-galactosidase, etc.

The terms Fv, Fc, Fd, Fab, or $F(ab')_2$ are used with their standard meanings (see, e.g., Harlow et al., Antibodies A Laboratory Manual, Cold Spring Harbor Laboratory, (1988)).

As used herein, an "engineered antibody" describes a type of altered antibody, i.e., a full-length synthetic antibody (e.g., a chimeric, humanized, reshaped or immunologically edited human antibody as opposed to an antibody fragment) in which a portion of the light and/or heavy chain variable domains of a selected acceptor antibody are replaced by analogous parts from one or more donor antibodies which have specificity for the selected epitope. For example, such molecules may include antibodies characterized by a humanized heavy chain associated with an unmodified light chain (or chimeric light chain), or vice versa. Engineered antibodies may also be characterized by alteration of the nucleic acid sequences encoding the acceptor antibody light and/or heavy variable domain framework regions in order to retain donor antibody binding specificity. These antibodies can comprise replacement of one or more CDRs (preferably all) from the acceptor antibody with CDRs from a donor antibody described herein.

A "chimeric antibody" refers to a type of engineered antibody which contains naturally-occurring variable region (light chain and heavy chains) derived from a donor antibody in association with light and heavy chain constant regions derived from an acceptor antibody from a heterologous species.

A "humanized antibody" refers to a type of engineered antibody having its CDRs derived from a non-human donor immunoglobulin, the remaining immunoglobulin-derived parts of the molecule being derived from one (or more) human immunoglobulin(s). In addition, framework support residues may be altered to preserve binding affinity (see, e.g., Queen et al., Proc. Nat'l. Acad. Sci. USA, 86:10029-10032 (1989), Hodgson et al., Bio/Technology, 9:421 (1991)).

An "immunologically edited antibody" refers to a type of engineered antibody in which changes are made in donor and/or acceptor sequences to edit regions in respect of cloning artifacts, germ line enhancements, etc. aimed at reducing the likelihood of an immunological response to the antibody on the part of a patient being treated with the edited antibody.

The term "donor antibody" refers to an antibody (monoclonal, or recombinant) which contributes the nucleic acid sequences of its variable regions, CDRs, or other functional fragments or analogs thereof to a first immunoglobulin partner, so as to provide the altered immunoglobulin coding region and resulting expressed altered antibody with the antigenic specificity and neutralizing activity characteristic of the donor antibody. One donor antibody suitable for use in this invention is a Fab fragment of a human neutralizing monoclonal antibody designated as Fab Hu19. Fab Hu19 is defined as a having the variable light chain DNA and amino acid sequences Hu 19A as shown in Figures 2, 3, 4A and 4B.

The term "acceptor antibody" refers to an antibody (monoclonal, or recombinant) from a source genetically unrelated to the donor antibody, which contributes all (or any portion, but preferably all) of the nucleic acid sequences encoding its heavy and/or light chain framework regions and/or its heavy and/or light chain constant regions to the first immunoglobulin partner. Preferably a human antibody is the acceptor antibody.

"CDRs" are defined as the complementarity determining region amino acid sequences of an antibody which are the hypervariable regions of immunoglobulin heavy and light chains. See, e.g., Kabat et al., Sequences of Proteins of Immunological Interest, 4th Ed., U.S. Department of Health and Human Services, National Institutes of Health (1987). There are three heavy chain and three light chain CDRs (or CDR regions) in the variable portion of an immunoglobulin. Thus, "CDRs" as used herein refers to all three heavy chain CDRs, or all three light chain CDRs (or both all heavy and all light chain CDRs, if appropriate). CDRs provide the majority of contact residues for the binding of the antibody to the antigen or epitope. CDRs of interest in this invention are derived from donor antibody variable heavy and light chain sequences, and include analogs of the naturally occurring CDRs, which analogs also share or retain the same antigen binding specificity and/or neutralizing ability as the donor antibody from which they were derived.

By "sharing the antigen binding specificity or neutralizing ability" is meant, for example, that although Fab Hu19 may be characterized by a certain level of antigen affinity, a CDR encoded by a nucleic acid sequence of Fab Hu19 in an appropriate structural environment may have a lower, or higher affinity. It is expected that CDRs of Fab Hu19 in such environments will nevertheless recognize the same epitope(s) as does the intact Fab Hu19. A "functional fragment" is a partial heavy or light chain variable sequence (e.g., minor deletions at the amino or carboxy terminus of the immunoglobulin variable region) which retains the same antigen binding specificity and/or neutralizing ability as the antibody from which the fragment was derived.

An "analog" is an amino acid sequence modified by at least one amino acid, wherein said modification can be chemical or a substitution or a rearrangement of a few amino acids (i.e., no more than 10), which modification permits the amino acid sequence to retain the biological characteristics, e.g., antigen specificity and high affinity, of the unmodified sequence. For example, (silent) mutations can be constructed, via substitutions, when certain endonuclease restriction sites are created within or surrounding CDR-encoding regions.

Analogs may also arise as allelic variations. An "allelic variation or modification" is an alteration in the nucleic acid sequence encoding the amino acid or peptide sequences of the invention. Such variations or modifications may be due to degeneracy in the genetic code or may be deliberately engineered to provide desired characteristics. These variations or modifications may or may not result in alterations in any encoded amino acid sequence.

The term "effector agents" refers to non-protein carrier molecules to which the altered antibodies, and/or natural or synthetic light or heavy chains of the donor antibody or other fragments of the donor antibody may be associated by conventional means. Such non-protein carriers can include conventional carriers used in the diagnostic field, e.g., polystyrene or other plastic beads, polysaccharides, e.g., as used in the BIAcore (Pharmacia) system, or other non-protein substances useful in the medical field and safe for administration to humans and animals. Other effector agents may include a macrocycle, for chelating a heavy metal atom, or radioisotopes. Such effector agents may also be useful to increase the half-life of the altered antibodies, e.g., polyethylene glycol.

II. Combinatorial Cloning:

As mentioned above, a number of problems have hampered the direct application of the hybridoma technology of G. Kohler and C. Milstein (Nature 256: 495-497 (1975)) to the generation and isolation of human monoclonal antibodies. Among these are a lack of suitable fusion partner myeloma cell lines used to form hybridoma cell lines as well as the poor stability of such hybridomas even when formed. These shortcomings are further exacerbated in the case of RSV because of the paucity of viral specific B cells in the perpherial circulation. Therefore, the molecular biological approach of combinatorial cloning is preferred.

Combinatorial cloning is disclosed generally in PCT Publication No. WO90/14430. Simply stated, the goal of combinatorial cloning is to transfer to a population of bacterial cells the immunological genetic capacity of a human cell, tissue or organ. It is preferred to employ cells, tissues or organs which are immunocompetent. Particularly useful sources include, without limitation, spleen,

thymus, lymph nodes, bone marrow, tonsil and perpherial blood lymphocytes. The cells may be optionally RSV stimulated *in vitro*, or selected from donors which are known to have produced an immune response or donors who are HIV^+ but asymptomatic.

The genetic information isolated from the donor cells can be in the form of DNA or RNA and is conveniently amplified by Polymerase Chain Reaction (PCR) or similar techniques. When isolated as RNA the genetic information is preferably converted into cDNA by reverse transcription prior to amplification. The amplification can be generalized or more specifically tailored. For example, by a careful selection of PCR primer sequences, selective amplification of immunoglobulin genes or subsets within that class of genes can be achieved.

Once the component gene sequences are obtained, in this case the genes encoding the variable regions of the various heavy and light antibody chains, the light and heavy chain genes are associated in random combinations to form a random combinatorial library. Various recombinant DNA vector systems have been described to facilitate combinatorial cloning (see: PCT Publication No. WO90/14430 supra, Scott and Smith, Science 249:386-406 (1990) or U. S. Patent 5,223,409). Having generated the combinatorial library, the products can, after expression, be conveniently screened by biopanning with RSV F protein or, if necessary, by epitope blocked biopanning as described in more detail below.

Initially it is generally preferred to use Fab fragments of mAbs for combinatorial cloning and screening and then to convert the Fabs to full length mAbs after selection of the desired candidate molecules. However, single chain antibodies can also be used for cloning and screening.

III. Antibody Fragments

The present invention contemplates the use of Fab fragments or $F(ab')_2$ fragments to derive full-length mAbs directed against the F protein of RSV. Although these fragments may be independently useful as protective and therapeutic agents *in vivo* against RSV-mediated conditions or *in vitro* as part of an RSV diagnostic, they are employed herein as a component of a reshaped human antibody.

A Fab fragment contains the entire light chain and amino terminal portion of the heavy chain; and an F(ab')$_2$ fragment is the fragment formed by two Fab fragments bound by additional disulfide bonds. RSV binding monoclonal antibodies provide sources of Fab fragments and F(ab')$_2$ fragments and can be obtained via combinatorial phage library (see, e.g., Winter et al., Ann. Rev. Immunol., 12:433-455 (1994) or Barbas et al. (Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992)) which are both hereby incorporated by reference in their entirety).

IV. Anti-RSV Antibody Amino Acid and Nucleotide Sequences of Interest

The Fab Hu19 or other antibodies described herein may contribute sequences, such as variable heavy and/or light chain peptide sequences, framework sequences, CDR sequences, functional fragments, and analogs thereof, and the nucleic acid sequences encoding them, useful in designing and obtaining various altered antibodies which are characterized by the antigen binding specificity of the donor antibody.

As one example, the present invention thus provides variable light chain and variable heavy chain sequences from the RSV human Fab Hu19A-D and sequences derived therefrom.

The nucleic acid sequences of this invention, or fragments thereof, encoding the variable light chain and heavy chain peptide sequences are also useful for mutagenic introduction of specific changes within the nucleic acid sequences encoding the CDRs or framework regions, and for incorporation of the resulting modified or fusion nucleic acid sequence into a plasmid for expression. For example, silent substitutions in the nucleotide sequence of the framework and CDR-encoding regions can be used to create restriction enzyme sites which would facilitate insertion of mutagenized CDR (and/or framework) regions. These CDR-encoding regions may be used in the construction of reshaped human antibodies of this invention.

Taking into account the degeneracy of the genetic code, various coding sequences may be constructed which encode the variable heavy and light chain amino acid sequences, and CDR sequences of the invention as well as functional

fragments and analogs thereof which share the antigen specificity of the donor antibody. The isolated nucleic acid sequences of this invention, or fragments thereof, encoding the variable chain peptide sequences or CDRs can be used to produce altered antibodies, e.g., chimeric or humanized antibodies, or other engineered antibodies of this invention when operatively combined with a second immunoglobulin partner.

It should be noted that in addition to isolated nucleic acid sequences encoding portions of the altered antibody and antibodies described herein, other such nucleic acid sequences are encompassed by the present invention, such as those complementary to the native CDR-encoding sequences or complementary to the modified human framework regions surrounding the CDR-encoding regions. Such sequences include all nucleic acid sequences which by virtue of the redundancy of the genetic code are capable of encoding the same amino acid sequence as given in Figures 2 and 3. Other useful DNA sequences encompassed by this invention include those sequences which hybridize under stringent hybridization conditions (See: T. Maniatis et al., Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), pages 387 to 389) to the DNA sequences encoding the antibodies of Figures 2 and 3 and which retain the antigen binding properties of those antibodies. An example of one such stringent hybridization condition is hybridization at 4XSSC at 65°C, followed by a washing in 0.1XSSC at 65°C for an hour. Alternatively an exemplary stringent hybridization condition is in 50% formamide, 4XSSC at 42°C. Preferably, these hybridizing DNA sequences are at least about 18 nucleotides in length, i.e., about the size of a CDR.

V. Altered Immunoglobulin Coding Regions and Altered Antibodies

Altered immunoglobulin coding regions encode altered antibodies which include engineered antibodies such as chimeric antibodies, humanized, reshaped and immunologically edited human antibodies. A desired altered immunoglobulin coding region contains CDR-encoding regions in the form of Fab regions that encode peptides having the antigen specificity of an RSV antibody, preferably a high

affinity antibody such as provided by the present invention, inserted into an acceptor immunoglobulin partner.

When the acceptor is an immunoglobulin partner, as defined above, it includes a sequence encoding a second antibody region of interest, for example an Fc region. Immunoglobulin partners may also include sequences encoding another immunoglobulin to which the light or heavy chain constant region is fused in frame or by means of a linker sequence. Engineered antibodies directed against functional fragments or analogs of RSV may be designed to elicit enhanced binding with the same antibody.

The immunoglobulin partner may also be associated with effector agents as defined above, including non-protein carrier molecules, to which the immunoglobulin partner may be operatively linked by conventional means.

Fusion or linkage between the immunoglobulin partners, e.g., antibody sequences, and the effector agent may be by any suitable means, e.g., by conventional covalent or ionic bonds, protein fusions, or hetero-bifunctional cross-linkers, e.g., carbodiimide, glutaraldehyde, and the like. Such techniques are known in the art and readily described in conventional chemistry and biochemistry texts.

Additionally, conventional linker sequences which simply provide for a desired amount of space between the second immunoglobulin partner and the effector agent may also be constructed into the altered immunoglobulin coding region. The design of such linkers is well known to those of skill in the art.

In addition, signal sequences for the molecules of the invention may be modified to enhance expression. For example the reshaped human antibody having the signal sequence and CDRs derived from the Fab Hu19 heavy chain sequence, may have the original signal peptide replaced with another signal sequence such as the Campath leader sequence (Page, M. J. et al., BioTechnology 9:64-68(1991)).

An exemplary altered antibody, a reshaped human antibody, contains a variable heavy and the entire light chain peptide or protein sequence having the antigen specificity of Fab Hu19, fused to the constant heavy regions C_{H-1}-C_{H-3} derived from a second human antibody.

In still a further embodiment, the engineered antibody of the invention may have attached to it an additional agent. For example, the procedure of recombinant DNA technology may be used to produce an engineered antibody of the invention in which the Fc fragment or CH2 CH3 domain of a complete antibody molecule has been replaced by an enzyme or other detectable molecule (i.e., a polypeptide effector or reporter molecule).

Another desirable protein of this invention may comprise a complete antibody molecule, having full length heavy and light chains, or any discrete fragment thereof, such as the Fab or $F(ab')_2$ fragments, a heavy chain dimer, or any minimal recombinant fragments thereof such as an F_v or a single-chain antibody (SCA) or any other molecule with the same specificity as the selected donor Fab Hu19. Such protein may be used in the form of an altered antibody, or may be used in its unfused form.

Whenever the immunoglobulin partner is derived from an antibody different from the donor antibody, e.g., any isotype or class of immunoglobulin framework or constant regions, an engineered antibody results. Engineered antibodies can comprise immunoglobulin (Ig) constant regions and variable framework regions from one source, e.g., the acceptor antibody, and one or more (preferably all) CDRs from the donor antibody, e.g., the anti-RSV antibody described herein. In addition, alterations, e.g., deletions, substitutions, or additions, of the acceptor mAb light and/or heavy variable domain framework region at the nucleic acid or amino acid levels, or the donor CDR regions may be made in order to retain donor antibody antigen binding specificity or to reduce potential immunogenicity. In the present invention, a preferred mutation is the alteration of the consensus N-linked glycosylation site in CDR2 of the Hu19A and Hu19B heavy chain, as exemplified in the heavy chains of Hu19C and Hu19D (Fig. 2) (SEQ ID NO'S 7 and 8).

Such engineered antibodies are designed to employ one (or both) of the variable heavy and/or light chains of the RSV mAb (optionally modified as described) or one or more of the below-identified heavy or light chain CDRs. The engineered antibodies of the invention are neutralizing, i.e., they desirably inhibit virus growth *in vitro* and *in vivo* in animal models of RSV infection.

Such engineered antibodies may include a reshaped human antibody containing the human heavy and light chain constant regions fused to the RSV antibody functional fragments. A suitable human (or other animal) acceptor antibody may be one selected from a conventional database, e.g., the KABAT® database, Los Alamos database, and Swiss Protein database, by homology to the nucleotide and amino acid sequences of the donor antibody. A human antibody characterized by a homology to the framework regions of the donor antibody (on an amino acid basis) may be suitable to provide a heavy chain constant region and/or a heavy chain variable framework region for insertion of the donor CDRs. A suitable acceptor antibody capable of donating light chain constant or variable framework regions may be selected in a similar manner. It should be noted that the acceptor antibody heavy and light chains are not required to originate from the same acceptor antibody.

Desirably the heterologous framework and constant regions are selected from human immunoglobulin classes and isotypes, such as IgG (subtypes 1 through 4), IgM, IgA and IgE. The Fc domains are not limited to native sequences, but include mutant variants known in the art that alter function. For example, mutations have been described in the Fc domains of certain IgG antibodies that reduce Fc-mediated complement and Fc receptor binding (see, e.g., A. R. Duncan et al., Nature 332:563-564 (1988); A. R. Duncan and G. Winter, Nature 332:738-740 (1988); M.-L. Alegre et al., J. Immunol. 148:3461-3468 (1992); M.-H. Tao et al., J. Exp. Med. 178:661-667 (1993); V. Xu et al. J. Biol. Chem., 269:3469-2374 (1994)), alter clearance rate (J.-K. Kim et al., Eur. J. Immunol. 24:542-548 (1994), and reduce structural heterogeneity (S. Angal et al., Mol. Immunol. 30:105-108 (1993)). Also, other modifications are possible such as oligomerization of the antibody by addition of the tailpiece segment of IgM and other mutations (R. I. F. Smith and S. L. Morrison, Biotechnology 12:683-688 (1994); R. I. F. Smith et al., J. Immunol., 154: 2226-2236 (1995)) or addition of the tailpiece segment of IgA (I. Kariv et al., J. Immunol. 157: 29-38 (1996). However, the acceptor antibody need not comprise only human immunoglobulin protein sequences. For instance a gene may be constructed in which a DNA sequence encoding part of a human immunoglobulin chain is fused to

a DNA sequence encoding a non-immunoglobulin amino acid sequence such as a polypeptide effector or reporter molecule.

The altered antibody thus preferably has the structure of a natural human antibody or a fragment thereof, and possesses the combination of properties required for effective therapeutic use, e.g., treatment of RSV mediated diseases in man, or for diagnostic uses.

It will be understood by those skilled in the art that an altered antibody may be further modified by changes in variable domain amino acids without necessarily affecting the specificity and high affinity of the donor antibody (i.e., an analog). It is anticipated that heavy and light chain amino acids may be substituted by other amino acids either in the variable domain frameworks or CDRs or both. Particularly preferred is the immunological editing of such reconstructed sequences as illustrated in the examples herein.

In addition, the variable or constant region may be altered to enhance or decrease selective properties of the molecules of the instant invention, as described above. For example, dimerization, binding to Fc receptors, or the ability to bind and activate complement (see, e.g., Angal et al., Mol. Immunol, 30:105-108 (1993), Xu et al., J. Biol. Chem, 269:3469-3474 (1994), Winter et al., EP 307,434-B).

Such antibodies are useful in the prevention and treatment of RSV mediated disorders, as discussed below.

VI. Production of Altered antibodies and Engineered Antibodies

The resulting reshaped human antibodies of this invention can be expressed in recombinant host cells, e.g., COS, CHO or myeloma cells. A conventional expression vector or recombinant plasmid is produced by placing these coding sequences for the altered antibody in operative association with conventional regulatory control sequences capable of controlling the replication and expression in, and/or secretion from, a host cell. Regulatory sequences include promoter sequences, e.g., CMV promoter, and signal sequences, which can be derived from other known antibodies. Similarly, a second expression vector can be produced having a DNA sequence which encodes a complementary antibody light or heavy

chain. Preferably this second expression vector is identical to the first except insofar as the coding sequences and selectable markers are concerned, so to ensure as far as possible that each polypeptide chain is functionally expressed. Alternatively, the heavy and light chain coding sequences for the altered antibody may reside on a single vector.

A selected host cell is co-transfected by conventional techniques with both the first and second vectors (or simply transfected by a single vector) to create the transfected host cell of the invention comprising both the recombinant or synthetic light and heavy chains. The transfected cell is then cultured by conventional techniques to produce the engineered antibody of the invention. The production of the antibody which includes the association of both the recombinant heavy chain and light chain is measured in the culture by an appropriate assay, such as ELISA or RIA. Similar conventional techniques may be employed to construct other altered antibodies and molecules of this invention.

Suitable vectors for the cloning and subcloning steps employed in the methods and construction of the compositions of this invention may be selected by one of skill in the art. For example, the conventional pUC series of cloning vectors, may be used. One vector used is pUC19, which is commercially available from supply houses, such as Amersham (Buckinghamshire, United Kingdom) or Pharmacia (Uppsala, Sweden). Additionally, any vector which is capable of replicating readily, has an abundance of cloning sites and selectable genes (e.g., antibiotic resistance), and is easily manipulated may be used for cloning. Thus, the selection of the cloning vector is not a limiting factor in this invention.

Similarly, the vectors employed for expression of the engineered antibodies according to this invention may be selected by one of skill in the art from any conventional vectors. Preferred vectors include for example plasmids pCD or pCN. The vectors also contain selected regulatory sequences (such as CMV promoters) which direct the replication and expression of heterologous DNA sequences in selected host cells. These vectors contain the above described DNA sequences which code for the engineered antibody or altered immunoglobulin coding region.

In addition, the vectors may incorporate the selected immunoglobulin sequences modified by the insertion of desirable restriction sites for ready manipulation.

The expression vectors may also be characterized by genes suitable for amplifying expression of the heterologous DNA sequences, e.g., the mammalian dihydrofolate reductase gene (DHFR). Other preferable vector sequences include a poly A signal sequence, such as from bovine growth hormone (BGH) and the betaglobin promoter sequence (betaglopro). The expression vectors useful herein may be synthesized by techniques well known to those skilled in this art.

The components of such vectors, e.g. replicons, selection genes, enhancers, promoters, signal sequences and the like, may be obtained from commercial or natural sources or synthesized by known procedures for use in directing the expression and/or secretion of the product of the recombinant DNA in a selected host. Other appropriate expression vectors of which numerous types are known in the art for mammalian, bacterial, insect, yeast, and fungal expression may also be selected for this purpose.

The present invention also encompasses a cell line transfected with a recombinant plasmid containing the coding sequences of the engineered antibodies or altered immunoglobulin molecules thereof. Host cells useful for the cloning and other manipulations of these cloning vectors are also conventional. However, most desirably, cells from various strains of *E. coli* are used for replication of the cloning vectors and other steps in the construction of altered antibodies of this invention.

Suitable host cells or cell lines for the expression of the engineered antibody or altered antibody of the invention are preferably mammalian cells such as CHO, COS, a fibroblast cell (e.g., 3T3), and myeloid cells, and more preferably a CHO or a myeloid cell. Human cells may be used, thus enabling the molecule to be modified with human glycosylation patterns. Alternatively, other eukaryotic cell lines may be employed. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art. See, e.g., Sambrook et al., cited above.

Bacterial cells may prove useful as host cells suitable for the expression of the recombinant Fabs of the present invention (see, e.g., Plückthun, A., Immunol.

Rev., 130:151-188 (1992)). The tendency of proteins expressed in bacterial cells to be in an unfolded or improperly folded form or in a non-glycosylated form does not pose as great a concern as Fabs are not normally glycosylated and can be engineered for exported expression thereby reducing the high concentration that facilitates misfolding. Nevertheless, any recombinant Fab produced in a bacterial cell would have to be screened for retention of antigen binding ability. If the molecule expressed by the bacterial cell was produced and exported in a properly folded form, that bacterial cell would be a desirable host. For example, various strains of *E. coli* used for expression are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*, *Streptomyces*, other bacilli and the like may also be employed in this method.

Where desired, strains of yeast cells known to those skilled in the art are also available as host cells, as well as insect cells, e.g. *Drosophila* and *Lepidoptera* and viral expression systems. See, e.g. Miller et al., Genetic Engineering, 8:277-298, Plenum Press (1986) and references cited therein.

The general methods by which the vectors of the invention may be constructed, the transfection methods required to produce the host cells of the invention, and culture methods necessary to produce the altered antibody of the invention from such host cell are all conventional techniques. Likewise, once produced, the altered antibodies of the invention may be purified from the cell culture contents according to standard procedures of the art, including ammonium sulfate precipitation, affinity columns, column chromatography, gel electrophoresis and the like. Such techniques are within the skill of the art and do not limit this invention.

Yet another method of expression of resphaped antibodies may utilize expression in a transgenic animal, such as described in U. S. Patent No. 4,873,316. This relates to an expression system using the animal's casein promoter which when transgenically incorporated into a mammal permits the female to produce the desired recombinant protein in its milk.

Once expressed by the desired method, the engineered antibody is then examined for *in vitro* activity by use of an appropriate assay. Presently conventional

ELISA assay formats are employed to assess qualitative and quantitative binding of the altered antibody to RSV. Additionally, other *in vitro* assays and *in vivo* animal models may also be used to verify neutralizing efficacy prior to subsequent human clinical studies performed to evaluate the persistence of the altered antibody in the body despite the usual clearance mechanisms.

VII. Therapeutic/Prophylactic Uses

This invention also relates to a method of treating humans experiencing RSV-related symptoms which comprises administering an effective dose of antibodies including one or more of the altered antibodies described herein or fragments thereof.

The therapeutic response induced by the use of the molecules of this invention is produced by the binding to RSV and thus subsequently blocking RSV propagation. Thus, the molecules of the present invention, when in preparations and formulations appropriate for therapeutic use, are highly desirable for those persons experiencing RSV infection. For example, longer treatments may be desirable when treating seasonal episodes or the like. The dose and duration of treatment relates to the relative duration of the molecules of the present invention in the human circulation, and can be adjusted by one of skill in the art depending upon the condition being treated and the general health of the patient.

The altered antibodies, antibodies and fragments thereof of this invention may also be used alone or in conjunction with other antibodies, particularly human or humanized mAbs reactive with other epitopes on the F protein or other RSV target antigens as prophylatic agents.

The mode of administration of the therapeutic and prophylatic agents of the invention may be any suitable route which delivers the agent to the host. The altered antibodies, antibodies, engineered antibodies, and fragments thereof, and pharmaceutical compositions of the invention are particularly useful for parenteral administration, i.e., subcutaneously, intramuscularly, intravenously, or intranasally.

Therapeutic and prophylacticagents of the invention may be prepared as pharmaceutical compositions containing an effective amount of the altered antibody

of the invention as an active ingredient in a pharmaceutically acceptable carrier. An aqueous suspension or solution containing the antibody, preferably buffered at physiological pH, in a form ready for injection is preferred. The compositions for parenteral administration will commonly comprise a solution of the engineered antibody of the invention or a cocktail thereof dissolved in an pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be employed, e.g., 0.4% saline, 0.3% glycine, and the like. These solutions are sterile and generally free of particulate matter. These solutions may be sterilized by conventional, well known sterilization techniques (e.g., filtration). The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, etc. The concentration of the antibody of the invention in such pharmaceutical formulation can vary widely, i.e., from less than about 0.5%, usually at or at least about 1% to as much as 15 or 20% by weight and will be selected primarily based on fluid volumes, viscosities, etc., according to the particular mode of administration selected.

Thus, a pharmaceutical composition of the invention for intramuscular injection could be prepared to contain 1 mL sterile buffered water, and between about 1 ng to about 100 mg, e.g. about 50 ng to about 80 mg or more preferably, about 5 mg to about 75 mg of an engineered antibody of the invention. Similarly, a pharmaceutical composition of the invention for intravenous infusion could be made up to contain about 250 ml of sterile Ringer's solution, and about 1 to about 75 and preferably 5 to about 50 mg/ml of an engineered antibody of the invention. Actual methods for preparing parenterally administrable compositions are well known or will be apparent to those skilled in the art and are described in more detail in, for example, Remington's Pharmaceutical Science, 15th ed., Mack Publishing Company, Easton, Pennsylvania.

It is preferred that the therapeutic and prophylactic agents of the invention, when in a pharmaceutical preparation, be present in unit dose forms. The appropriate therapeutically effective dose can be determined readily by those of skill in the art. To effectively treat an inflammatory disorder in a human or other animal,

one dose of approximately 0.1 mg to approximately 20 mg per 70 kg body weight of a protein or an antibody of this invention should be administered parenterally, preferably i.v. or i.m. (intramuscularly). Such dose may, if necessary, be repeated at appropriate time intervals selected as appropriate by a physician.

The altered antibodies and engineered antibodies of this invention may also be used in diagnostic regimens, such as for the determination of RSV mediated disorders or tracking progress of treatment of such disorders. As diagnostic reagents, these altered antibodies may be conventionally labeled for use in ELISAs and other conventional assay formats for the measurement of RSV levels in serum, plasma or other appropriate tissue, or the release by human cells in culture. The nature of the assay in which the altered antibodies are used are conventional and do not limit this disclosure.

The antibodies, altered antibodies or fragments thereof described herein can be lyophilized for storage and reconstituted in a suitable carrier prior to use. This technique has been shown to be effective with conventional immunoglobulins and art-known lyophilization and reconstitution techniques can be employed.

The following examples illustrate various aspects of this invention including the construction of exemplary engineered antibodies and expression thereof in suitable vectors and host cells, and are not to be construed as limiting the scope of this invention. All amino acids are identified by conventional three letter or single letter codes. All necessary restriction enzymes, plasmids, and other reagents and materials were obtained from commercial sources unless otherwise indicated. All general cloning ligation and other recombinant DNA methodology were as performed in T. Maniatis et al., cited above, or the second edition thereof (1989), eds. Sambrook et al., by the same publisher ("Sambrook et al.").

Example A

Conversion of Hu19 Fab to mAb Version A: Direct Cloning

For expression in mammalian cells, the heavy chain variable region and the light chain variable and constant regions from the Fab clone 19 plasmid(C. Barbas III et al., Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992) and PCT application

Publication No. WO 94/06448, Application No. PCT/US93/08786 Cell Line Clone 19 referenced therein as ATCC Accession No. 69072) herein designated Hu19 Fab, were cloned into derivatives of plasmid PCDN (Nambi, A. et al., Molecular and Cellular Biochemistry 131:75-86 (1994), in which the expression of the antibody chain is driven by the CMV promoter. Plasmid PCD-HC68B is used for cloning and expressing full length heavy chains and plasmid PCN-HuLC, for cloning and expressing full length light chains (Figure 1 shows the strategy for cloning of version A of the Hu19 mAb).

In the initial constructs, the changes in the sequence at the amino terminus, introduced by the PCR primers used for cloning, were not altered. For the heavy chain, the variable region was extracted from the Hu19 Fab plasmid (C. Barbas III et al., Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992)) as an *Xho*1-*Bsp*120I fragment and introduced into the same sites in PCD-HC68B. The *Xho*I site was introduced at the amino terminus by the PCR primer and, when cloned into PCD-HC68B at the same site is preceded in frame by the Campath leader sequence (Page, J.M. et al., Biotechnology 9:64-68 (1991). The *Bsp*120I site is a naturally occurring, highly conserved sequence at the beginning of the CH1 domain, and when cloned into PCD-HC68B at the same site is in frame with the remaining sequence for the CH1 through CH3 regions of human IgG1 (Figure 1). In the resulting construct, Hu19AHcpcd, the amino acids immediately following the Campath leader are EVQLLEE (Fig. 2 SEQ ID NO 5, AMINO ACIDS 20 - 26), where the residues LE are encoded by the nucleotide sequence for the XhoI cloning site. The complete nucleotide sequence for the plasmid Hu19AHcpcd is shown in Fig. 4A (SEQ ID NO 14).

Of note, sequence analysis revealed base differences from the published sequence (C. Barbas III et al., Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992), PCT publication WO94/06448) within the heavy chain region from the Hu19 Fab plasmid. The changes encode amino acid differences at positions 15 and 16 (14 and 15 according to consensus numbering of Kabat et al (Sequences of Proteins of Immunological Interest, fifth edition, NIH Publication No. 91-3242, 1991):

PG in the Hu19AHcpcd vector versus LA in the published sequence (see Fig 2 of this application and Fig. 4 of WO94/06448)). This discrepancy must represent an error in the original published sequence. PG at these positions is the consensus sequence in the closest homologues among published human antibodies (Kabat et al., Sequences of Proteins of Immunological Interest, fifth edition, NIH Publication No. 91-3242, 1991) and in the likely germline parent sequence (see below, version B). In addition, sequences derived from 3 independent clonings initiated with the Hu19 Fab plasmid all encoded PG at these positions.

For the light chain, the variable and constant regions of the Hu19 Fab plasmid were cloned as a *Sac*1/*Xba*1 fragment into the same sites in the pCN-HuLcvector. Both restriction sites correspond to restriction sites introduced by the primers used in the PCR amplification. The Sac1 site is introduced at the amino terminus by the PCR primer and, when cloned into pCN-HuLC at the same site, is preceded in frame by the Campath leader sequence (Page, J.M. et al., Biotechnology 9:64-68 (1991). The first 2 amino acids of the mature light chain are therefore deleted. In the resulting construct, Hu19ALcpcn, the first 2 amino acids immediately following the leader are EL (Fig. 3, part A), where the residues EL are encoded by the nucleotide sequence for the *Sac*1 cloning site . The PCR primer used at the carboxy terminus of the constant region introduces a nucleotide substitution which changes the amino acid at position 202 of the mature light chain, from a serine to a leucine (Fig 3, part B). The *Xba*1 restriction site, introduced by the same PCR primer, lies outside the coding region and has no effect on the final amino acid sequence of the mature light chain. The complete nucleotide sequence of the plasmid Hu19Apcn is shown in Fig. 4B.

As for the heavy chain above, there was a sequence discrepancy for the light chain between the published sequence (C. Barbas III et al., Proc. Nat'l. Acad. Sci. (USA) 89: 10164-10168 (1992), PCT publication WO94/06448) and the sequence obtained in the Hu19ALcpcn vector. A single base change resulted in glycine in Hu19ALcpcn in place of glutamic acid at position 97 (also consensus position 97 in Kabat et al (Sequences of Proteins of Immunological Interest, fifth edition, NIH Publication No. 91-3242, 1991)) in framework 4 (see Fig. 3 of this application and

Fig. 4 of WO94/06448). Glycine, but not glutamic acid, is encoded at this position in a human J germline J mini-gene and glutamic acid was not observed among a large collection of human antibody sequences (Kabat et al., "Sequences of Proteins of Immunological Interest", fifth edition, NIH Publication No. 91-3242, 1991). Also as for the heavy chain, the glycine encoding sequence was observed for 3 separate clonings from the original Fab 19 vector (Barbas et al., Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992), PCT publication WO94/06448). These results demonstrate that the originally published sequence for Fab 19 light chain is in error.

The Hu19AHcpcd and Hu19ALcpcn set of vectors were used to produce antibody Hu19A in COS cells and in CHO cells.

Example B

Version B :Cloning Of The Edited Fab Hu19 Heavy and Light Chains

In cloning the variable region of the Fab 19 heavy chain, non-consensus amino acid changes relative to the predicted germline sequence were introduced at the amino terminus by the PCR primer (C. Barbas III et al., Proc. Nat'l. Acad. Sci. (USA) 89, 10164-10168 (1992)). To determine the likely amino terminus of the heavy chain, the peptide sequence of the variable region of the Fab 19 heavy chain was aligned with all known human germline heavy chain sequences. Based on the results of this alignment, the germline amino terminus is predicted to be either QVQLVE or EVQLVE rather than the sequence EVQLLEE present in version A. To correct the N-terminus, the original Fab clone 19 heavy chain peptide was aligned with human heavy chain sequences previously cloned at SmithKline Beecham. A clone designated 97B27, which was obtained via PCR amplification from the beginning of its leader sequence, had the acceptable N-terminus of QVQLVE and was used to replace this region in the Fab19 heavy chain. Specifically, the Fab19 heavy chain in the Hu19 Fab plasmid was PCR amplified using a constant region primer which spanned the naturally occurring *Bsp120 I* site at the beginning of CH1, and a variable region primer which created a *Pvu*II site (corresponding to the site naturally occurring in clone 97B27) at amino acids 3 and 4 of the mature protein. This primer also introduced changes in the coding sequence at the amino terminus of the Fab19 heavy chain, coding for the amino acid sequence of QLVE for amino acids 3-6 instead of QLLEE, as in the version A construct. The PCR fragment was cut with restriction enzymes *Pvu*II and *Bsp* 120I, and, through a series of cloning steps, was combined with 97B27 at its *Pvu*II site. The resulting clone, designated Hu19BHcpcd, contained the leader and first 3 amino acids of the variable region of clone 97B27 and coded for the consensus sequence QVQLVE at its amino terminus (Fig. 2). The nucleotide sequence of Hu19BHcpcd is shown in Fig. 4C (SEQ ID NO: 20) for the region encoding the heavy chain. Sequences differing from Hu19AHcpcd are bolded.

In cloning the variable region of the Fab clone 19 light chain, changes were introduced at the amino terminus for cloning purposes, by the PCR primer, such that the first 4 amino acids of the Fab19 light chain are EIEL. To determine the likely amino terminus of the light chain, the peptide sequence of the variable region of the Fab19 light chain was aligned with all known human germline kappa chain sequences. Based on the results of this alignment, the germline amino terminus is predicted to be DIQM. To convert the amino terminus of Fab19 Lc to the predicted germline sequence, Fab19 light chain was aligned with human kappa chain sequences previously cloned at SB. A Clone designated AG1-37, which is the kappa chain obtained from cell line AG1-37 obtained by PCR amplification from the middle of its leader sequence, had the desired N-terminus and was used to introduce the corrections into the Fab19 light chain. The N-terminal portion of the leader sequence was provided by the expression vector and was the consensus sequence for this family of leader regions. For this construct, the light chain coding region was excised from the Hu19 Fab vector (Fig. 1) as a *HinfI/Xba* I fragment. *Hinf*I recognizes a site which spans amino acids 18 an 19 of the mature protein and is also present in clone AG1-37. Through a series of cloning steps, the *Hinf*I/*Xba*I fragment of the Fab19 light chain was ligated to the *Hinf*I site in clone AG1-37. The final construct consisted of the leader and first 18 amino acids of the AG1-37 variable region linked to the variable and constant regions of the Fab 19 light chain, beginning at amino acid 19 of the V-region. The resulting clone, designated Hu19BLcpcn, is altered only in the region encoding the first four amino acids of the variable region, coding for the consensus sequence DIQM (SEQ ID NO: 11, AMINO ACIDS 21 - 24) instead of EIEL present in version A (Fig. 3A). The nucleotide sequence for plasmid Hu19BLcpcn is shown in Fig. 4D (SEQ ID NO: 22) for the region encoding the light chain. Sequences differing from Hu19ALcpcn are bolded.

The vector set of Hu19BHcpcd and Hu19BLcpcn was used to produce antibody Hu19B in COS cells and in CHO cells.

Example C

Versions C & D: Mutation Of CDR2 Of Hu19B Heavy Chain To Eliminate a Glycosylation Site

An N-linked glycosylation site is encoded within the CDR2 loop of the heavy chain. This glycosylation adds the potential for heterogeneity in the mAb produced in eucaryotic cells and for interference in binding antigen. To eliminate this glycosylation site, mutations were introduced separately at two different residues via PCR overlap technology. For the first mutation the serine at position 61 of the mature Hu19B heavy chain was substituted with alanine, to create Hu19C heavy chain. For the second substitution, the asparagine at position 59 was changed to glutamine, to create Hu19D heavy chain.

```
SITGGSNGINYADSVKR S61A Substitution (SEQ ID NO: 1)
|||||||||||:|||||
SITGGSNGINYSDSVKR Original  HuB CDR2 (SEQ ID NO: 2)
|||||||||:|||||||
SITGGSNGIQYSDSVKR N59Q Substitution (SEQ ID NO: 3)
```

Specifically, the mutations were introduced via the PCR overlap technique using one set of primers encoding the mutation and a second set of primers annealing to sequences within the CMV promoter and the CH2 constant region in plasmid Hu19Bpcd, as the outside 5' and 3' primers, respectfully. The final PCR product was digested with restriction enzymes, EcoR1 and *Bsp*120I, and cloned into the Hu19BHcpcd vector at the same sites to create Hu19CHcpcd (Ser to Ala mutation) and Hu19DLcpcd (Asp to Gln mutation) (Fig. 2). The final constructs were sequenced to verify that the mutations were present. The nucleotide sequences of the heavy chain regions in Hu19CHcpcd and Hu19DHcpcd are shown in Figs. 4E and 4F (SEQ ID NO'S 25 AND 27). Differences from Hu19Hcpcd are bolded.

Example D

Version C :Cloning Of The Edited Constant Region

In the original cloning the of the Fab19 light chain, a change was purposely introduced at the caboxy terminus by the PCR primer to eliminate a naturally occurring *Sac*1 site (Barbas et al, supra). Consequently, the amino acid at position 202 of the Fab19 light chain was changed from a serine to a leucine. This change was corrected as follows. Plasmid Hu19BLcpcn was cut with *Eco*R1 and *Bbs*1, a naturally occurring restriction site near the amino terminus of human kappa constant region and a 405 bp fragment, containing the nucleotide sequence coding for the leader, variable region, and first 5 amino acids of the kappa constant region, was isolated. Plasmid Lcvector4, a puc18 derivative containing a normal human kappa constant region with a *Xba*I site just distal to the coding region, was cut with *Bbs*1 and *Xba*1 and a 321 bp fragment containing the nucleotide sequence coding for the entire kappa constant region beginning at amino acid 6 was isolated. This fragment contains the naturally occurring *Sac*1 site near the end of the carboxy terminus and codes for a serine at position 202. Plasmid Hu19BLcpcn was also cut with *Eco*R1 and *Xba*1 and a 4947 bp fragment, containing the remainder of the vector sequence from plasmid Hu19BLcpcn, was isolated. The three fragments were ligated together to create Hu19CLcpcn. The amino acid sequence of the Hu19C light chain is shown in Figs. 3A and 3B (SEQ ID NO'S 11 and 12) and the nucleotide sequence of the light chain region is shown in Fig 4G (SEQ ID NO: 29). Differences from Hu19ALcpcn are bolded. The vector Hu19CLcpcn, was used with Hu19CHcpcd or Hu19DHcpcd to produce antibody Hu19C and Hu19D, respectively, in COS cells and in CHO cells.

Example E

Production of Hu19 Mabs in mammalian cells:

For initial characterization, the mAb constructs for each version, Hu19A heavy and light chain, Hu19B heavy and light chain, Hu19C heavy and light chain, and Hu19D heavy with Hu19C light chain, were expressed in COS cells essentially as described in Current Protocols in Molecular Biology (edited by F. M. Ausubel et

al. 1988. John Wiley & Sons, vol. 1, section 9.1). On day 1 after the transfection, the culture growth medium was replaced with a serum-free medium which was changed on day 3. The serum-free medium was a proprietary formulation but satisfactory results are obtained using DMEM supplemented with ITS^{TM} Premix (insulin, transferrin, selenium mixture - Collaborative Research, Bedford, MA) and 1 mg/ml BSA. The mAb was prepared from the day 3 + day 5 conditioned medium by standard protein A affinity chromatography methods (e.g., as described in Protocols in Molecular Biology) using, for example, Prosep A affinity resin (Bioprocessing Ltd., UK).

To produce larger quantities of the Hu19 mAb (100-200 mgs), the vectors were introduced into a proprietary CHO cell system. However, similar results will be obtained using dhfr$^-$ CHO cells as previously described (P. Hensley et al., J. Biological Chemistry 269:23949-23958 (1994)). Briefly, a total of 30ug of linearized plasmid DNA (15ug each of the A, B, C or D/C set of heavy chain and light chain vectors) was electroporated into $1x10^7$ cells. The cells were initially selected in nucleoside-free medium in 96 well plates. After three to four weeks, media from growth positive wells was screened for human immunoglobulin using an ELISA assay. The highest expressing colonies were expanded and selected in increasing concentrations of methotrexate for amplification of the transfected vectors. The antibody was purified from conditioned medium by standard procedures using protein A affinity chromatography (Protein A sepharose, Pharmacia) followed by size exclusion chromatography (Superdex 200, Pharmacia).

The concentration and the antigen binding activity of the eluted antibody are measured by ELISA. The antibody containing fractions are pooled and further purified by size exclusion chromatography. As expected for any such antibody, by SDS-PAGE, the predominant protein product migrated at approximately 150 kDa under non-reducing conditions and as 2 bands of 50 and 25 kDa under reducing conditions. For antibody produced in CHO cells, the purity was > 90%, as judged by SDS-PAGE, and the concentration was accurately determined by amino acid analysis.

Example F

Preparation of Fab from Hu19B mAb: Samples with and without glycosylation in heavy chain CDR2

Purification of MAbs

Each mAb was purified using an essentially similar purification procedure that is detailed here for mAb 19B. Conditioned media (2L) from a 6 day culture was harvested, sterile filtered and applied to a 2.5 X 5.1cm Protein A (Pharmacia, fast flow) equilibrated in 20mM sodium phosphate, 150mM sodium chloride, pH 7 (PBS) at a linear flow rate of 98cm/h. The column was washed with equilibration buffer and eluted with 100mM glycine pH 2.5. Elution fractions containing the mAb were immediately adjusted to pH 5.0 with 1M sodium hydroxide and applied at a concentration of 4.2 mg/mL to a Superdex 200 size exclusion column (2.6 x 70 cm) equilibrated in 20 mM sodium phosphate buffer containing 150 mM NaCl, pH7.0. Monomeric mAb that was retained by the column at an apparent molecular weight of around 150 kDa was pooled and concentrated by ultrafiltration to 5mg/mL, and stored at 4°C after sterile filtration.

Electrophoretic analysis of MAb19B and MAb 19C

By reduced SDS-PAGE, mAb 19B resolved as 2 major bands at 52 kDa and 28kDa corresponding to the heavy and light chains of IgG respectively, with an additional band at 59 kDa representing about 7% of the total protein (Fig. 5). LC/mass spectrometry analysis of the two heavy chains following excision from an SDS-PAGE and proteolytic digestion (see below), confirmed that the 59 kDa species represented an additional glycoform of mAb 19B that contained carbohydrate at the predicted V_H glycosylation site. In contrast, reduced SDS-PAGE analysis of mAb 19C (Fig. 5), in which this V_H glycosylation site is removed, showed that this mAb contains only the lower molecular weight (52 kDa) heavy chain species, as expected.

Carbohydrate Analysis of mAb 19B

The Hu19B construct contains an additional consensus sequence for N-linked glycosylation in the variable region of the heavy chain, -Asn^{59}-Tyr-Ser-, in addition to the normal glycosylation site in the C_H2 domain of the heavy chain, -Asn^{299}-Ser-Thr-. Analysis of both heavy chain bands by liquid chromatography, electrospray mass spectrometry (LC-ELMS) following reduction, alkylation, and tryptic digestion revealed that the 59 kDa band contains a variant that is glycosylated at Asn^{59} in addition to being glycosylated at Asn^{299}. The carbohydrate at Asn^{59} is predominantly biantennary, core fucosylated carbohydrates having two sialic acid residues. This is a common carbohydrate structure found in CHO-expressed glycoproteins (such as sCR-1 and sCD4), but it differs from the carbohydrate found at the Asn299 site which lacks sialic acid altogether.

Purification of mAb 19B Glycovariant

Mab 19B (2 mg) was dialyzed against 20 mM Tris, pH 8.5 and applied to a 0.5 x 5cm Mono Q column (Pharmacia) equilibrated in the same buffer at a linear flow rate of 300cm/h. The column was washed with equilibration buffer and eluted with a 20 column volume gradient from 0 mM to 50 mM NaCl in the same buffer (Fig. 6). Fractions containing the glycovariant, as determined by SDS-PAGE, were pooled, dialyzed against PBS, sterile filtered and stored at 4°C.

Preparation of Fabs by Proteolytic Digestion

mAb19B (48mg) was removed and the pH adjusted to 7.0 with dilute sodium hydroxide. 2.5ml of 100mM sodium phosphate buffer containing 10mM EDTA, pH 7.3; 1.3ml of 100mM cysteine in 10mM sodium phosphate buffer containing 1mM EDTA; and 20ul of crystalline papain (Boehringer, 10mg/ml) were added. The sample was incubated at 37°C for 20h and the digest applied to a 1.5 x 2.6cm Protein G column equilibrated in PBS at a linear flow rate of 67cm/h. The column was washed with PBS and the nonbound fraction containing the Fab was collected and concentrated to 5ml in an Amicon ultrafiltration cell fitted with a 10,000 molecular weight cut-off membrane and applied to a 2.6 X 70cm Superdex 200 (Pharmacia) size exclusion column equilibrated in PBS at a linear flow rate of 23cm/h. FAb (total yield, 12mg) eluted as a monomer on the size exclusion column and analysis by non-reduced SDS-PAGE revealed a major band at 45kDa and the glycoform at 47kDa.

Separation of Fab Glycovariant

The mixture of glycosylated and unglycosylated Fab from cleaved mAb 19B was dialyzed against 20 mM sodium acetate, pH 4.5 and applied (4mg) to a 0.5 X 5cm Mono S column (Pharmacia) at 300cm/h equilibrated with 20mM sodium acetate buffer, pH 4.5. The column was then washed with equilibration buffer and eluted isocratically with the equilibration buffer containing 100 mM NaCl. Glycosylated Fab eluted after 5 column volumes whereas the unglycosylated FAb was retained longer, eluting after 6 column volumes. Fractions that contained only glycosylated Fab, as judged by SDS-PAGE, were pooled, diluted 1:1 with starting buffer and reapplied to a 0.16 X 5cm Mono S column at 300cm/h. The Fab was once again eluted with 100 mM NaCl and fractions most enriched for glycosylated Fab were pooled, dialyzed against PBS pH 7.0, and sterile filtered. By SDS-PAGE analysis this fraction was enriched >90% with the glycosylated species (Fig. 7). The process yielded 3.3 mg of unglycosylated Fab and 0.16 mg of glycosylated Fab, respectively.

Example G

Binding of Hu19 mAb and Fab clone 19 proteins to recombinant F protein

Binding of the various antibody constructs to recombinant F protein was measured in a standard solid phase ELISA. Antigen diluted in PBS pH 7.0 was adsorbed onto polystyrene round-bottom microplates (Dynatech, Immunolon II) for 18 hours. Wells were then aspirated and blocked with 0.5% boiled casein (BC) in PBS containing 1% Tween 20 (PBS/0.05% BC) for 2 hours. Antibodies (50 μl/well) were diluted to varying concentrations in PBS/0.5% BC containing 0.025% Tween 20 and incubated in antigen coated wells for one hour. Plates were washed three times with PBS containing 0.05% Tween 20, using a Titertek 320 microplate washer, followed by addition of HRP-labelled protein A/G (50 μl) diluted 1:5000. After washing three times, TMBlue substrate (TSI, #TM102) was added and plates were incubated an additional 15 minutes. The reaction was stopped by addition of 1 NH_2SO_4 and absorbance read at 450 nm using a Biotek ELISA reader.

The antigen binding epitope of Fab19 and mAb construct 19B were examined in a competition ELISA. The test antibody construct was mixed with increasing concentrations of RSMU19 or B4 and added to F protein-coated wells. The epitope regions recognized by mAbs RSMU19 and B4 have been previously described in Arbiza et al., J. Gen'l Virol. 73:2225-34 (1992). The concentration of Fab19 or mAb 19B used in competition studies was determined previously to give 90% maximal binding to F antigen. Binding of Fab19 or mAb 19B in the presence of other mAbs was detected using HRP-labelled goat anti-human IgG. The reaction was developed as stated above.

Fab19 and amAb constructs 19A or 19B, demonstrated equivalent binding to rF protein based on molar concentrations. Binding of Fab19 or mAb 19B to rF (recombinant F) protein was inhibited by mAb B4 but not by RSMU19 indicating that the epitope region recognized by these constructs is localized to region aa 255-275 of the F protein (Table 1).

Table 1: Viral F Protein Epitope Recognized by mAb 19B

Construct	Binding to rF	Competition Binding to rF	
	EC_{50} (M)	RSMU19 mAb (aa 429)*	B4 mAb (aa 268, 272, 275)
Fab19	10^{-9}	-	+
mAb 19A	10^{-9}	not tested	not tested
mAb 19B	10^{-9}	-	+

* amino acid residues critical for antigen recognition

The mAb 19B also showed specific binding to RSV infected cells indicating recognition of the F protein as displayed in its native form. VERO cells infected with approximately 50 TCID50 RS Long virus were fixed in 90% methanol when CPE reached > 90% and were used as antigen in the ELISA format described above. Binding of biotinylated mAb 19B was detected with HRP-labelled -Streptavidin. In this assay, the EC50 for mAb 19B was 34 +/- ng/ml.

Example H

<u>In vitro antiviral activity of the Hu19 Antibodies</u>

The ability of Fab fragments to inhibit virus-induced cell fusion was determined using a modification of the in vitro microneutralization assay described by Beeler et al (J. of Virology 63: 2941-2950 (1989)). In this assay, 50 ul of RS Long strain virus (approximately 100 $TCID_{50}$/well; American Type Culture Collection ATCC VR-26) were mixed with 0.1 ml VERO cells (5 x 10^3/well; ATCC CCL-81) in Minimum Essential Media (MEM) containing 2% FCS, for 4 hours at 37°C, 5% CO_2. Serial two-fold dilution (in duplicate) of test samples (50 ul) were then added to wells containing virus-infected cells. Control cultures contained cells incubated with virus only (positive virus control) or cells incubated with media alone. Cultures were incubated at 37°C in 5% CO_2 for 6 days at which time cytopathic effects (CPE) in virus control wells were ≥ 90%. Neutralization assays

were performed as described above except that serial dilutions of test samples were mixed with 100 $TCID_{50}$ of RS virus (50 ul each) for 2 hours at 37°C in 5% CO_2 before the addition of VERO cells (5 x 10^3).

Microscopic examination for cytopathic effects were confrimed by ELISA. Media was aspirated from cultures and replaced with 50 ul of 90% methanol/0.6% H_2O_2. After 10 minutes, fixative was aspirated and plates were air dried overnight. Viral antigen was detected in the fixed cultures using biotinylated human/bovine chimeric derivative of mAb B4 (RSCHB4; 1 ug/ml), followed by HRP-labelled streptavidin (Boehringer-Mannheim) diluted 1:10,000 (each lot was titrated to determine the optimal concentration). The reaction was developed using TMBlue and stopped by addition of 1 N H_2SO_4. Absorbance was measured at 450nm ($O.D._{450}$).

Fusion-inhibition or neutralization titers were defined as the reciprocal dilution of test sample, or concentration of antibody, which caused a 50% reduction in ELISA signal (ED_{50}) as compared to virus controls. Based on the curve generated in the ELISA by the standard virus titration, a 50% reduction in $O.D._{450}$ in wells corresponded to ≥ 90% reduction in virus titer. To determine the ED_{50}, mean absorbance for replicate cultures (per dilution of test sample) was plotted against dilution of sample. Calculation of the 50% point, defined as (mean absorbance virus-infected cells + mean absorbance uninfected cells)/2, was based on regression analysis of the dose titration.

SB 209763 is a humanized derivative of RSMU19 as described in P. R. Tempest et al., Biotechnology 9, 266-271 (1991). To determine the effects of coadministration of mAb19B and SB 209763 on *in vitro* fusion-inhibition, the antibodies were titrated alone and in combination. Antibody interactions were analyzed using MacSynergy ™ II software.

The *in vitro* antiviral titers of the mAb constructs generated either by direct cloning (version A) or after introduction of various sequence modfications (versions B-D) demonstrated potent neutralization and fusion-inhibition activity against a prototype RSV Long strain (Table 2). mAb 19B was also shown to neutralize clinical isolates representing multiple antigenic variants of RSV collected over the

1993/1994 season in the Philadelphia PA area (Table 3). When mAb 19B was co-administered with a second antibody directed to a different F protein epitope (SB 209763, critical residue aa 429), the effect on inhibition of virus growth in infected cell cultures was additive (data not shown).

The antiviral titers of the mAb constructs were approximately 5 to 10-fold lower than the titers obtained with the corresponding Fab constructs - Fab19, Fab19A or B (Table 2). Fab19 is the original Fab protein produced directly in *E. coli* from the clone 19 plasmid,whereas Fab19A and Fab19B were derived by papain cleavage from the coresponding full length mAbs. Removal of the N-linked glycoslation site encoded within the CDR2 loop of the heavy chain by cloning had no effect on the overall antiviral activity of the mAb (Table 2; construct C compared to A and B). In addition, enrichment of the mAb19B construct for normally glycosylated antibody did not significantly alter the *in vitro* fusion-inhibition titer (Table 4). However, enrichment for the glycovariant Fab fragment resulted in a 2 to 10-fold reduction in *in vitro* antiviral activity compared to normally glycosylated Fab fragment (Table 4).

Table 2: Antiviral Activity of 19A, 19B, 19C, and 19D Constructs Against RS Long strain virus

Construct	Neutralization EC_{50}, ug/ml	Fusion-Inhibition EC_{50}	
		(ug/ml)	(nM)
Fab19	0.34 ± 0.25*	0.22	4.4
Fab19A	not tested	0.16	3
Fab19B	not tested	0.12 ± 0.06	2.4
mAb 19A	2.2	2.8 ± 1.9	18.9
mAb 19B	1.5	2.3 ± 1.9	15.3
mAb 19C	not tested	2.4	16
mAb 19D	not tested	2.6	17.3

* mean ± standard deviation

Table 3: Fusion-Inhibition Activity of mAb 19B against Clinical Isolates of RSV

Virus Isolate	Fusion-Inhibition Titer EC$_{50}$ (ug/ml)	
	mAb19B	SB 209763
RS Long (prototype A1)	2.3 ± 1.9	1.3 ± 0.8
RS 9320 (prototype B1)	0.59	2.5 + 1.1
A1 - V1763	2.79	1.95
A2 - 847	0.89	0.27
A2 - 626	0.35	0.36
A3 - 7062	2.65	1.67
A4 - 6652	2.1	1.52
B1 - 6973	1.77	2.22
B2 - 6556	1.49	2.05
B3 - 447	1.78	1.7

Table 4: Antiviral Activity of 19B Glycovariants

Construct	% glycovariant*	Fusion-Inhibition Titer EC_{50} (ug/ml)
mAb 19B	40%	2.5 ± 1.5
Fraction A	< 5%	1.8 ± 0.8
Fraction B	40%	3.8 ± 0.9
Fab19B	< 10%	0.12 ± 0.06
Fraction A	1%	0.89
Fraction B	94%	1.5 ± 0.2
Fraction C	99%	3.7

* mAb or Fab fragments were untreated or run on MonoQ (Mab) or MonoS (Fab) columns to separate glycosylated versus minimally glycosylated forms in the varible region.

Example I

In vivo Activity of mAb 19B; Prophylaxis and Therapy in Balb/c Mouse Model.

Balb/c mice (5/group) were inoculated intraperitoneally with doses ranging from 0.06 mg/kg to 5 mg/kg of mAb 19B either 24 hours prior (prophylaxis) or 4 days after (therapy) intranasal infection with 10^5 PFU of the A2 strain of human RSV. Mice were sacrificed 5 days after infection. Sera was obtained to determine antibody levels and lungs were homogenized to determine virus titers. Virus was undetectable in the lungs of mice treated prophylactically with ≥ 1.25 ug mAb 19B, and corresponding serum concentrations of ≥ 5 ug/ml (Table 5). Higher doses of mAb 19B were required for complete viral clearance when mAb was administered therapeutically (5 mg/kg).

Table 5: mAb 19B Prophylaxis and Therapy in Balb/c Mice

		Prophylaxis		Therapy	
Treatment	**Dose (mg/kg**	**Lung Virus Titer (log_{10}/g lung)**	**Serum Concentration (ug/ml)**	**Lung Virus Titer (log_{10}/g lung)**	**Serum Concentration (ug/ml)**
mAb 19B	5	$\leq$1.7	15.6	$\leq$1.7	13.2
	1.25	$\leq$1.7	5.0	2.5 $\pm$ 0.4	2.1
	0.31	3.2 $\pm$ 0.3	0.79	3.8 $\pm$ 0.2	0.61
	0.06	3.8 $\pm$ 0.6	0.17	4.5 $\pm$ 0.1	0.08
PBS	-	5.2 $\pm$ 0.1	$\leq$0.02	4.7 $\pm$ 0.3	$\leq$0.036

The results of examples G through I establish that the Hu19 antibodies have potent antiviral activity in vitro against a broad range of native RSV isolates of both type A and B, and show prophylactic and therapeutic efficacy in vivo in animal models. Thus, the Hu19 antibodies, most preferably Hu19C or Hu19D, are candidates for therapeutic, prophylactic, and diagnostic application in man.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: [illegible] Keith [illegible]
[illegible], Susan [illegible]
[illegible], [illegible]
Sweet, Raymond A.

(ii) TITLE OF THE INVENTION: Human Monoclonal Antibodies

(iii) NUMBER OF SEQUENCES: [illegible]

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SmithKline Beecham Corporation
(B) STREET: 709 Swedeland Road
(C) CITY: King of Prussia
(D) STATE: PA
(E) COUNTRY: U.S.A.
(F) ZIP: 19046

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: [illegible],149
(B) FILING DATE: 01-NOV-[illegible]

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Geiger, Kathleen
(B) REGISTRATION NUMBER: 35,880
(C) REFERENCE/DOCKET NUMBER: P50504

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 610-270-5968
(B) TELEFAX: 610-270-5090
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Ser Ile Thr Gly Gly Ser Asn Gly Ile Asn Tyr Ala Asp Ser Val Lys
 1               5                  10                  15
Arg
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Ser Ile Thr Gly Gly Ser Asn Gly Ile Asn Tyr Ser Asp Ser Val Lys
 1               5                  10                  15
Arg
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Ser Ile Thr Gly Gly Ser Asn Gly Ile Gln Tyr Ser Asp Ser Val Lys
 1               5                  10                  15
Arg
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1               5                  10                  15
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
            20                  25                  30
Glu Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
        35                  40                  45
Ser Tyr Ile Ser Ser Ser Gly Ser Thr Ile Tyr Tyr Ala Asp Ser Val
    50                  55                  60
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
65                  70                  75                  80
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
                85                  90                  95
Ala Arg
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1               5                  10                  15
Val His Ser Glu Val Gln Leu Leu Glu Val Glu Ser Gly Gly Gly Leu
            20                  25                  30
```

Arg Ile Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Thr Thr
35 40 45

Leu Ser Gly Tyr Thr Met His Trp Val Arg Gln Ala Pro Gly Lys Gly
50 55 60

Leu Glu Trp Val Ser Ser Ile Thr Gly Gly Ser Asn Thr Ile Asn Tyr
65 70 75 80

Ser Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys
85 90 95

Asn Ser Leu Tyr Leu Gln Met Asn Ser Leu Thr Ala Glu Asp Thr Ala
100 105 110

Val Tyr Tyr Cys Ala Thr Ala Pro Ile Ala Pro Pro Tyr Phe Asp His
115 120 125

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
130 135

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Arg Gly
1 5 10 15

Val Gln Cys Gln Val Gln Leu Val Val Glu Ser Gly Gly Gly Leu Arg
20 25 30

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Thr Thr Leu
35 40 45

Ser Gly Tyr Thr Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
50 55 60

Glu Trp Val Ser Ser Ile Thr Gly Gly Ser Asn Phe Ile Asn Tyr Ser
65 70 75 80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn
85 90 95

Ser Leu Tyr Leu Gln Met Asn Ser Leu Thr Ala Glu Asp Thr Ala Val
100 105 110

Tyr Tyr Cys Ala Thr Ala Pro Ile Ala Pro Pro Tyr Phe Asp His Trp
115 120 125

Gly Gln Gly Thr Leu Val Thr Val Ser Ser
130 135

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Arg Gly
1 5 10 15

Val Gln Cys Gln Val Gln Leu Val Val Glu Ser Gly Gly Gly Leu Arg
20 25 30

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Thr Thr Leu
35 40 45

Ser Gly Tyr Thr Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
50 55 60

Glu Trp Val Ser Ser Ile Thr Gly Gly Ser Asn Phe Ile Asn Tyr Ala
65 70 75 80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn
85 90 95

Ser Leu Tyr Leu Gln Met Asn Ser Leu Thr Ala Glu Asp Thr Ala Val

100 105 110

Tyr Tyr Cys Ala Thr Ala Pro Ile Ala Pro Pro Tyr Phe Asp His Trp

115 120 125

Gly Gln Gly Thr Leu Val Thr Val Ser Ser

130 135

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Arg Gly

1 5 10 15

Val Gln Cys Gln Val Gln Leu Val Val Glu Ser Gly Gly Gly Leu Arg

20 25 30

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Thr Thr Leu

35 40 45

Ser Gly Tyr Thr Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu

50 55 60

Glu Trp Val Ser Ser Ile Thr Gly Gly Ser Asn Phe Ile Gln Tyr Ser

65 70 75 80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn

85 90 95

Ser Leu Tyr Leu Gln Met Asn Ser Leu Thr Ala Glu Asp Thr Ala Val

100 105 110

Tyr Tyr Cys Ala Thr Ala Pro Ile Ala Pro Pro Tyr Phe Asp His Trp

115 120 125

Gly Gln Gly Thr Leu Val Thr Val Ser Ser
130 135

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15
Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
20 25 30
Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45
Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60
Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80
Glu Asp Phe Ala Thr Tyr Tyr Cys
85

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
1 5 10 15

Val His Ser Glu Leu Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val
20 25 30

Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Thr Gln Ser Val Ser Asn
35 40 45

Phe Leu Asn Trp Tyr Gln Gln Lys Pro Gly Glu Ala Pro Thr Leu Leu
50 55 60

Ile Tyr Asp Ala Ser Thr Ser Gln Ser Gly Val Pro Ser Arg Phe Ser
65 70 75 80

Gly Ser Gly Ser Gly Met Asp Phe Ser Leu Thr Ile Ser Ser Leu Gln
85 90 95

Pro Glu Asp Leu Ala Met Tyr Tyr Cys Gln Ala Ser Ile Asn Thr Pro
100 105 110

Leu Phe Gly Gly Gly Thr Arg Ile Asp Met Arg Arg
115 120

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu Trp Leu Arg
1 5 10 15

Gly Ala Arg Cys Asp Ile Gln Met Asn Phe Leu Asn Trp Tyr Gln Gln
20 25 30
Lys Pro Gly Glu Ala Pro Thr Leu Leu Ile Tyr Asp Ala Ser Thr Ser
35 40 45
Gln Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Met Asp
50 55 60
Phe Ser Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Leu Ala Met Tyr
65 70 75 80
Tyr Cys Gln Ala Ser Ile Asn Thr Pro Leu Phe Gly Gly Gly Thr Arg
85 90 95
Ile Asp Met Arg Arg
100

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 106 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
1 5 10 15
Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
20 25 30
Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
35 40 45
Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
50 55 60
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
65 70 75 80

[illegible] Tyr Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys

[illegible]

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
1 5 10 15

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
20 25 30

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
35 40 45

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
50 55 60

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
65 70 75 80

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Leu Pro
85 90 95

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
100 105

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6284 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG     60
AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC    120
GGAGAATGGG CGGAACTGGG CGGAGTTAGG GGCGGGATGG GCGGAGTTAG GGGCGGGACT    180
ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG    240
GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT    300
GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAT TAATTCCCGG    360
GGATCGATCC GTCGACGTAC GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT    420
CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA    480
CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA    540
ATAGGGACTT TCCATTGACG TCAATGGGTG GACTATTTAC GGTAAACTGC CCACTTGGCA    600
GTACATCAAG TGTATCATAT GCCAAGTACG CCCCCTATTG ACGTCAATGA CGGTAAATGG    660
CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC    720
TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT    780
GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT CAATGGGAGT    840
TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC GTAACAACTC CGCCCCATTG    900
ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA TAAGCAGAGC TGGGTACGTG    960
AACCGTCAGA TCGCCTGGAG ACGCCATCGA ATTCTGAGCA CACAGGACCT CACCATGGGA   1020
TGGAGCTGTA TCATCCTCTT CTTGGTAGCA ACAGCTACAG GTGTCCACTC CGAGGTCCAA   1080
CTGCTCGAGG AGTCTGGGGG AGGCCTGGTC AGGCCTGGCG GGTCCCTAAG ACTCTCGTGT   1140
GCAGCCTCTG GAACCACCCT CAGTGGCTAT ACCATGCACT GGGTCCGCCA GGCTCCAGGG   1200
AAGGGGCTGG AGTGGGTCTC ATCCATTACT GGAGGTAGCA ACTTCATAAA CTACTCAGAC   1260
TCAGTGAAGG GCCGATTCAC CATCTCCAGA GACAACGCCA AGAACTCACT TTATCTGCAA   1320
ATGAACAGCC TGACAGCCGA GGACACGGCT GTCTATTATT GTGCGACCGC CCCTATAGCA   1380
CCGCCCTACT TTGACCACTG GGGCCAGGGA ACCCTGGTCA CCGTCTCCTC AGCCTCCACC   1440
AAGGGCCCAT CGGTCTTCCC CCTGGCACCC TCCTCCAAGA GCACCTCTGG GGGCACAGCG   1500
GCCCTGGGCT GCCTGGTCAA GGACTACTTC CCCGAACCGG TGACCGTGTC GTGGAACTCA   1560
GGCGCCCTGA CCAGCGGCGT GCACACCTTC CCGGCTGTCC TACAGTCCTC AGGACTCTAC   1620
TCCCTCAGCA GCGTGGTGAC TGTGCCCTCC AGCAGCTTGG GCACCCAGAC CTACATCTGC   1680
AACGTGAATC ACAAGCCCAG CAACACCAAG GTGGACAAGA AAGTTGAGCC CAAATCTTGT   1740
```

GAGAAAACT[illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
[illegible] [illegible] CAAGGACA[illegible] [illegible] [illegible] [illegible]
[illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
[illegible] [illegible] [illegible] [illegible] [illegible] CAGCACGTA[illegible]
CGTGTGGTCA GCGTCCTCAC CGTCCTGCAC CAGGACTGGC TGAATGGCAA GGAGTACAAG
TGCAAGGTCT CCAACAAAGC CCTCCCAGCC CCCATCGAGA AAACCATCTC CAAAGCCAAA
GGGCAGCCCC GAGAACCACA GGTGTACACC CTGCCCCCAT CCCGGGATGA GCTGACCAAG
AACCAGGTCA GCCTGACCTG CCTGGTCAAA GGCTTCTATC CCAGCGACAT CGCCGTGGAG
TGGGAGAGCA ATGGGCAGCC GGAGAACAAC TACAAGACCA CGCCTCCCGT GCTGGACTCC
GACGGCTCCT TCTTCCTCTA CAGCAAGCTC ACCGTGGACA AGAGCAGGTG GCAGCAGGGG
AACGTCTTCT CATGCTCCGT GATGCATGAG GCTCTGCACA ACCACTACAC GCAGAAGAGC
CTCTCCCTGT CTCCGGGTAA ATGATAGATA TCTACGTATG ATCAGCCTCG ACTGTGCCTT
CTAGTTGCCA GCCATCTGTT GTTTGCCCCT CCCCCGTGCC TTCCTTGACC CTGGAAGGTG
CCACTCCCAC TGTCCTTTCC TAATAAAATG AGGAAATTGC ATCGCATTGT CTGAGTAGGT
GTCATTCTAT TCTGGGGGGT GGGGTGGGGC AGGACAGCAA GGGGGAGGAT TGGGAAGACA
ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATGGAACC AGCTGGGGCT CGACAGCGCT
GGATCTCCCG ATCCCCAGCT TTGCTTCTCA ATTTCTTATT TGCATAATGA GAAAAAAAGG
AAAATTAATT TTAACACCAA TTCAGTAGTT GATTGAGCAA ATGCGTTGCC AAAAAGGATG
CTTTAGAGAC AGTGTTCTCT GCACAGATAA GGACAAACAT TATTCAGAGG GAGTACCCAG
AGCTGAGACT CCTAAGCCAG TGAGTGGCAC AGCATTCTAG GGAGAAATAT GCTTGTCATC
ACCGAAGCCT GATTCCGTAG AGCCACACCT TGGTAAGGGC CAATCTGCTC ACACAGGATA
GAGAGGGCAG GAGCCAGGGC AGAGCATATA AGGTGAGGTA GGATCAGTTG CTCCTCACAT
TTGCTTCTGA CATAGTTGTG TTGGGAGCTT GGATAGCTTG GACAGCTCAG GGCTGCGATT
TCGCGCCAAA CTTGACGGCA ATCCTAGCGT GAAGGCTGGT AGGATTTTAT CCCCGCTGCC
ATCATGGTTC GACCATTGAA CTGCATCGTC GCCGTGTCCC AAAATATGGG GATTGGCAAG
AACGGAGACC TACCCTGGCC TCCGCTCAGG AACGAGTTCA AGTACTTCCA AAGAATGACC
ACAACCTCTT CAGTGGAAGG TAAACAGAAT CTGGTGATTA TGGGTAGGAA AACCTGGTTC
TCCATTCCTG AGAAGAATCG ACCTTTAAAG GACAGAATTA ATATAGTTCT CAGTAGAGAA
CTCAAAGAAC CACCACGAGG AGCTCATTTT CTTGCCAAAA GTTTGGATGA TGCCTTAAGA
CTTATTGAAC AACCGGAATT GGCAAGTAAA GTAGACATGG TTTGGATAGT CGGAGGCAGT
TCTGTTTACC AGGAAGCCAT GAATCAACCA GGCCACCTTA GACTCTTTGT GACAAGGATC
ATGCAGGAAT TTGAAAGTGA CACGTTTTTC CCAGAAATTG ATTTGGGGAA ATATAAACTT
CTCCCAGAAT ACCCAGGCGT CCTCTCTGAG GTCCAGGAGG AAAAAGGCAT CAAGTATAAG
TTTGAAGTCT ACGAGAAGAA AGACTAACAG GAAGATGCTT TCAAGTTCTC TGCTCCCCTC
CTAAAGCTAT GCATTTTTAT AAGACCATGG GACTTTTGCT GGCTTTAGAT CAGCCTCGAC
TGTGCCTTCT AGTTGCCAGC CATCTGTTGT TTGCCCCTCC CCCGTGCCTT CCTTGACCCT

GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT CGCATTGTCT 3960
GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG 4020
GGAAGACAAT AGCAGGCATG CTGGGGATGC GGTGGGCTCT ATGGAACCAG CTGGGGCTCG 4080
ATCGAGTGTA TGACTGCGGC CGCGATCCCG TCGAGAGCTT GGCGTAATCA TGGTCATAGC 4140
TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA GCCGGAAGCA 4200
TAAAGTGTAA AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAATT GCGTTGCGCT 4260
CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT GCATTAATGA ATCGGCCAAC 4320
GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC 4380
TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT 4440
TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG 4500
CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG 4560
AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT 4620
ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA 4680
CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT 4740
GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC 4800
CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA 4860
GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG 4920
TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG 4980
TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT 5040
GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA 5100
CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC 5160
AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA 5220
CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA 5280
CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT 5340
TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT 5400
TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT 5460
TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT 5520
CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA 5580
ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG 5640
GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT 5700
TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG 5760
CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG 5820
TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC 5880
GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA 5940
CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC 6000
CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT 6060

[illegible]

[illegible]

[illegible]

[illegible]

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: [illegible] amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1               5                  10                  15
Val His Ser Glu Val Gln Leu Leu Glu Val
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Ser Pro Gly Lys
 1
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1               5                  10                  15
Val His Ser Glu Leu Thr Gln Ser Pro
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5681 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG    60
AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC   120
GGAGAATGGG CGGAACTGGG CGGAGTTAGG GGCGGGATGG GCGGAGTTAG GGGCGGGACT   180
ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG   240
GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT   300
GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAT TAATTCCCGG   360
GGATCGATCC GTCGACGTAC GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT   420
```

[illegible] [illegible]
[illegible] [illegible]
ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC CCACTTGGCA [illegible]
GTACATCAAG TGTATCATAT GCCAAGTACG CCCCCTATTG ACGTCAATGA CGGTAAATGG [illegible]
[illegible] ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC [illegible]
TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT [illegible]
GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT CAATGGGAGT 840
TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC GTAACAACTC CGCCCCATTG 900
ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA TAAGCAGAGC TGGGTACGTG 960
AACCGTCAGA TCGCCTGGAG ACGCCATCGA ATTCTGAGCA CACAGGACCT CACCATGGGA 1020
TGGAGCTGTA TCATCCTCTT CTTGGTAGCA ACAGCTACAG GTGTCCACTC CGAGCTCACC 1080
CAGTCTCCAT CCTCCCTGTC TGCATCTGTA GGAGACAGAG TCACCATCAC TTGCCGGGCA 1140
ACTCAGAGTG TTAGTAACTT TTTAAATTGG TATCAGCAGA AGCCAGGGGA AGCCCCTACG 1200
CTCCTGATCT ATGATGCATC CACTTCGCAA AGTGGGGTCC CATCAAGGTT CAGTGGCAGT 1260
GGATCTGGGA TGGATTTCAG TCTCACCATC AGCAGTCTGC AGCCTGAAGA TCTTGCAATA 1320
TATTACTGTC AAGCGAGTAT CAATACCCCG CTTTTCGGCG GAGGGACCAG AATAGATATG 1380
AGACGAACTG TGGCTGCACC ATCTGTCTTC ATCTTCCCGC CATCTGATGA GCAGTTGAAA 1440
TCTGGAACTG CCTCTGTTGT GTGCCTGCTG AATAACTTCT ATCCCAGAGA GGCCAAAGTA 1500
CAGTGGAAGG TGGATAACGC CCTCCAATCG GGTAACTCCC AGGAGAGTGT CACAGAGCAG 1560
GACAGCAAGG ACAGCACCTA CAGCCTCAGC AGCACCCTGA CGCTGAGCAA AGCAGACTAC 1620
GAGAAACACA AAGTCTACGC CTGCGAAGTC ACCCATCAGG GCCTGAGCTT GCCCGTCACA 1680
AAGAGCTTCA ACAGGGGAGA GTGTTAGTGA GATGATCCTC TAGAGTCATC TACGTATGAT 1740
CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT TTGCCCCTCC CCCGTGCCTT 1800
CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT 1860
CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG 1920
GGGAGGATTG GGAAGACAAT AGCAGGCATG CTGGGGATGC GGTGGGCTCT ATGGAACCAG 1980
CTGGGGCTCG ACAGCTCGAG CTAGCTTTGC TTCTCAATTT CTTATTTGCA TAATGAGAAA 2040
AAAAGGAAAA TTAATTTTAA CACCAATTCA GTAGTTGATT GAGCAAATGC GTTGCCAAAA 2100
AGGATGCTTT AGAGACAGTG TTCTCTGCAC AGATAAGGAC AAACATTATT CAGAGGGAGT 2160
ACCCAGAGCT GAGACTCCTA AGCCAGTGAG TGGCACAGCA TTCTAGGGAG AAATATGCTT 2220
GTCATCACCG AAGCCTGATT CCGTAGAGCC ACACCTTGGT AAGGGCCAAT CTGCTCACAC 2280
AGGATAGAGA GGGCAGGAGC CAGGGCAGAG CATATAAGGT GAGGTAGGAT CAGTTGCTCC 2340
TCACATTTGC TTCTGACATA GTTGTGTTGG GAGCTTGGAT CGATCCACCA TGGTTGAACA 2400
AGATGGATTG CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCG GCTATGACTG 2460
GGCACAACAG ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG CGCAGGGGCG 2520
CCCGGTTCTT TTTGTCAAGA CCGACCTGTC CGGTGCCCTG AATGAACTGC AGGACGAGGC 2580

AGCGCGGCTA TCGTGGCTGG CCACGACGGG CGTTCCTTGC GCAGCTGTGC TCGACGTTGT 2640
CACTGAAGCG GGAAGGGACT GGCTGCTATT GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC 2700
ATCTCACCTT GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA 2760
TACGCTTGAT CCGGCTACCT GCCCATTCGA CCACCAAGCG AAACATCGCA TCGAGCGAGC 2820
ACGTACTCGG ATGGAAGCCG GTCTTGTCGA TCAGGATGAT CTGGACGAAG AGCATCAGGG 2880
GCTCGCGCCA GCCGAACTGT TCGCCAGGCT CAAGGCGCGC ATGCCCGACG GCGAGGATCT 2940
CGTCGTGACC CATGGCGATG CCTGCTTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC 3000
TGGATTCATC GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC 3060
TACCCGTGAT ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC TCGTGCTTTA 3120
CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG ACGAGTTCTT 3180
CTGAGCGGGA CTCTGGGGTT CGAAATGACC GACCAAGCGA CGCCCAACCT GCCATCACGA 3240
GATTTCGATT CCACCGCCGC CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC 3300
GCCGGCTGGA TGATCCTCCA GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC 3360
TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT 3420
AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT 3480
CATGTCTGGA TCGCGGCCGC GATCCCGTCG AGAGCTTGGC GTAATCATGG TCATAGCTGT 3540
TTCCTGTGTG AAATTGTTAT CCGCTCACAA TTCCACACAA CATACGAGCC GGAAGCATAA 3600
AGTGTAAAGC CTGGGGTGCC TAATGAGTGA GCTAACTCAC ATTAATTGCG TTGCGCTCAC 3660
TGCCCGCTTT CCAGTCGGGA AACCTGTCGT GCCAGCTGCA TTAATGAATC GGCCAACGCG 3720
CGGGGAGAGG CGGTTTGCGT ATTGGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC 3780
GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT 3840
CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA 3900
GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC 3960
ATCACAAAAA TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC 4020
AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG 4080
GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC TCACGCTGTA 4140
GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG 4200
TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC 4260
ACGACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG 4320
GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGGACAGTAT 4380
TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT 4440
CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC 4500
GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT 4560
GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT 4620
AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT 4680
GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC 4740

[illegible]

TGTTATCACT CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA 5220

GATGCTTTTC TGTGACTGGT GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGT 5280

GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT 5340

TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC 5400

TGTTGAGATC CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA 5460

CTTTCACCAG CGTTTCTGGG TGAGCAAAAA CAGGAAGGCA AAATGCCGCA AAAAAGGGAA 5520

TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT TATTGAAGCA 5580

TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC 5640

AAATAGGGGT TCCGCGCACA TTTCCCCGAA AAGTGCCACC T 5681

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Leu Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
1 5 10

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1427 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
GAATTCGGTA CCATGGAGTT TGGGCTGAGC TGGGTTTTCC TCGTGGCTCT TTTAAGAGGT     60
GTCCAGTGTC AGGTGCAGCT GGTGGAGTCT GGGGGAGGCC TGGTCAGGCC TGGCGGGTCC    120
CTAAGACTCT CGTGTGCAGC CTCTGGAACC ACCCTCAGTG GCTATACCAT GCACTGGGTC    180
CGCCAGGCTC CAGGGAAGGG GCTGGAGTGG GTCTCATCCA TTACTGGAGG TAGCAACTTC    240
ATAAACTACT CAGACTCAGT GAAGGGCCGA TTCACCATCT CCAGAGACAA CGCCAAGAAC    300
TCACTTTATC TGCAAATGAA CAGCCTGACA GCCGAGGACA CGGCTGTCTA TTATTGTGCG    360
ACCGCCCCTA TAGCACCGCC CTACTTTGAC CACTGGGGCC AGGGAACCCT GGTCACCGTC    420
TCCTCAGCCT CCACCAAGGG CCCATCGGTC TTCCCCCTGG CACCCTCCTC CAAGAGCACC    480
TCTGGGGGCA CAGCGGCCCT GGGCTGCCTG GTCAAGGACT ACTTCCCCGA ACCGGTGACC    540
GTGTCGTGGA ACTCAGGCGC CCTGACCAGC GGCGTGCACA CCTTCCCGGC TGTCCTACAG    600
TCCTCAGGAC TCTACTCCCT CAGCAGCGTG GTGACCGTGC CCTCCAGCAG CTTGGGCACC    660
CAGACCTACA TCTGCAACGT GAATCACAAG CCCAGCAACA CCAAGGTGGA CAAGAAAGTT    720
GAGCCCAAAT CTTGTGACAA AACTCACACA TGCCCACCGT GCCCAGCACC TGAACTCCTG    780
GGGGGACCGT CAGTCTTCCT CTTCCCCCCA AAACCCAAGG ACACCCTCAT GATCTCCCGG    840
ACCCCTGAGG TCACATGCGT GGTGGTGGAC GTGAGCCACG AAGACCCTGA GGTCAAGTTC    900
AACTGGTACG TGGACGGCGT GGAGGTGCAT AATGCCAAGA CAAAGCCGCG GGAGGAGCAG    960
TACAACAGCA CGTACCGGGT GGTCAGCGTC CTCACCGTCC TGCACCAGGA CTGGCTGAAT   1020
GGCAAGGAGT ACAAGTGCAA GGTCTCCAAC AAAGCCCTCC CAGCCCCCAT CGAGAAAACC   1080
ATCTCCAAAG CCAAAGGGCA GCCCCGAGAA CCACAGGTGT ACACCCTGCC CCCATCCCGG   1140
GATGAGCTGA CCAAGAACCA GGTCAGCCTG ACCTGCCTGG TCAAAGGCTT CTATCCCAGC   1200
GACATCGCCG TGGAGTGGGA GAGCAATGGG CAGCCGGAGA ACAACTACAA GACCACGCCT   1260
CCCGTGCTGG ACTCCGACGG CTCCTTCTTC CTCTACAGCA AGCTCACCGT GGACAAGAGC   1320
AGGTGGCAGC AGGGGAACGT CTTCTCATGC TCCGTGATGC ATGAGGCTCT GCACAACCAC   1380
TACACGCAGA AGAGCCTCTC CCTGTCTCCG GGTAAATGAT AGATATC                 1427
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Arg Gly
 1               5                  10                  15
Val Gln Cys Gln Val Gln Leu Val
            20
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 732 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
GAATTCCATG GACATGAGGG TCCCCGCTCA GCTCCTAGGG CTCCTGCTGC TCTGGCTCCG    60
AGGTGCCAGA TGTGACATCC AGATGACCCA GTCTCCATCC TCCCTGTCTG CATCTGTAGG   120
AGACAGAGTC ACCATCACTT GCCGGGCAAG TCAGAGTGTT AGTAACTTTT TAAATTGGTA   180
TCAGCAGAAG CCAGGGGAAG CCCCTACGCT CCTGATCTAT GATGCATCCA CTTCGCAAAG   240
TGGGGTCCCA TCAAGGTTCA GTGGCAGTGG ATCTGGGATG GATTTCAGTC TCACCATCAG   300
CAGTCTGCAG CCTGAAGATC TTGCAATGTA TTACTGTCAA GCGAGTATCA ATACCCCGCT   360
TTTCGGCGGA GGGACCAGAA TAGATATGAG AGGAACTGTG GCTGCACCAT CTGTCTTCAT   420
CTTCCCGCCA TCTGATGAGC AGTTGAAATC TGGAACTGCC TCTGTTGTGT GCCTGCTGAA   480
TAACTTCTAT CCCAGAGAGG CCAAAGTACA GTGGAAGGTG GATAACGCCC TCCAATCGGG   540
TAACTCCCAG GAGAGTGTCA CAGAGCAGGA CAGCAAGGAC AGCACCTACA GCCTCAGCAG   600
CACCCTGACG CTGAGCAAAG CAGACTACGA GAAACACAAA GTCTACGCCT GCGAAGTCAC   660
CCATCAGGGC CTGAGCTTGC CCGTCACAAA GAGCTTCAAC AGGGGAGAGT GTTAGTGAGA   720
```

TGATCCTCTA GA 732

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Met Arg Val Pro Ala Gln Leu Leu Gly Leu Leu Leu Leu Trp Leu Arg
 1               5                  10                  15
Gly Ala Arg Cys Asp Ile Gln Met Thr
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
Leu Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
 1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1427 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
GAATTCGGTA CCATGGAGTT TGGGCTGAGC TGGGTTTTCC TCGTGGCTCT TTTAAGAGGT
GTCCAGTGTC AGGTGCAGCT GGTGGAGTCT GGGGGAGGCC TGGTCAGGCC TGGCGGGTCC
CTAAGACTCT CGTGTGCAGC CTCTGGAACC ACCCTCAGTG GCTATACCAT GCACTGGGTC
CGCCAGGCTC CAGGGAAGGG GCTGGAGTGG GTCTCATCCA TTACTGGAGG TAGCAACTTC
ATAAACTACG CAGACTCAGT GAAGGGCCGA TTCACCATCT CCAGAGACAA CGCCAAGAAC
TCACTTTATC TGCAAATGAA CAGCCTGACA GCCGAGGACA CGGCTGTCTA TTATTGTGCG
ACCGCCCCTA TAGCACCGCC CTACTTTGAC CACTGGGGCC AGGGAACCCT GGTCACCGTC
TCCTCAGCCT CCACCAAGGG CCCATCGGTC TTCCCCCTGG CACCCTCCTC CAAGAGCACC
TCTGGGGGCA CAGCGGCCCT GGGCTGCCTG GTCAAGGACT ACTTCCCCGA ACCGGTGACC
GTGTCGTGGA ACTCAGGCGC CCTGACCAGC GGCGTGCACA CCTTCCCGGC TGTCCTACAG
TCCTCAGGAC TCTACTCCCT CAGCAGCGTG GTGACCGTGC CCTCCAGCAG CTTGGGCACC
CAGACCTACA TCTGCAACGT GAATCACAAG CCCAGCAACA CCAAGGTGGA CAAGAAAGTT
GAGCCCAAAT CTTGTGACAA AACTCACACA TGCCCACCGT GCCCAGCACC TGAACTCCTG
GGGGGACCGT CAGTCTTCCT CTTCCCCCCA AAACCCAAGG ACACCCTCAT GATCTCCCGG
ACCCCTGAGG TCACATGCGT GGTGGTGGAC GTGAGCCACG AAGACCCTGA GGTCAAGTTC
AACTGGTACG TGGACGGCGT GGAGGTGCAT AATGCCAAGA CAAAGCCGCG GGAGGAGCAG
TACAACAGCA CGTACCGGGT GGTCAGCGTC CTCACCGTCC TGCACCAGGA CTGGCTGAAT
GGCAAGGAGT ACAAGTGCAA GGTCTCCAAC AAAGCCCTCC CAGCCCCCAT CGAGAAAACC
ATCTCCAAAG CCAAAGGGCA GCCCCGAGAA CCACAGGTGT ACACCCTGCC CCCATCCCGG
GATGAGCTGA CCAAGAACCA GGTCAGCCTG ACCTGCCTGG TCAAAGGCTT CTATCCCAGC
GACATCGCCG TGGAGTGGGA GAGCAATGGG CAGCCGGAGA ACAACTACAA GACCACGCCT
CCCGTGCTGG ACTCCGACGG CTCCTTCTTC CTCTACAGCA AGCTCACCGT GGACAAGAGC
AGGTGGCAGC AGGGGAACGT CTTCTCATGC TCCGTGATGC ATGAGGCTCT GCACAACCAC
TACACGCAGA AGAGCCTCTC CCTGTCTCCG GGTAAATGAT AGATATC
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ser Asn Phe Ile Asn Tyr Ala
1 5

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1427 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
GAATTCGGTA CCATGGAGTT TGGGCTGAGC TGGGTTTTCC TCGTGGCTCT TTTAAGAGGT    60
GTCCAGTGTC AGGTGCAGCT GGTGGAGTCT GGGGGAGGCC TGGTCAGGCC TGGCGGGTCC   120
CTAAGACTCT CGTGTGCAGC CTCTGGAACC ACCCTCAGTG GCTATACCAT GCACTGGGTC   180
CGCCAGGCTC CAGGGAAGGG GCTGGAGTGG GTCTCATCCA TTACTGGAGG TAGCAACTTC   240
ATACAATACT CAGACTCAGT GAAGGGCCGA TTCACCATCT CCAGAGACAA CGCCAAGAAC   300
TCACTTTATC TGCAAATGAA CAGCCTGACA GCCGAGGACA CGGCTGTCTA TTATTGTGCG   360
ACCGCCCCTA TAGCACCGCC CTACTTTGAC CACTGGGGCC AGGGAACCCT GGTCACCGTC   420
TCCTCAGCCT CCACCAAGGG CCCATCGGTC TTCCCCCTGG CACCCTCCTC CAAGAGCACC   480
TCTGGGGGCA CAGCGGCCCT GGGCTGCCTG GTCAAGGACT ACTTCCCCGA ACCGGTGACC   540
GTGTCGTGGA ACTCAGGCGC CCTGACCAGC GGCGTGCACA CCTTCCCGGC TGTCCTACAG   600
TCCTCAGGAC TCTACTCCCT CAGCAGCGTG GTGACCGTGC CCTCCAGCAG CTTGGGCACC   660
CAGACCTACA TCTGCAACGT GAATCACAAG CCCAGCAACA CCAAGGTGGA CAAGAAAGTT   720
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

ATCTTTAAAG CCAAAGGGCA GCCTCGAGAA CCACAGGTGT ACACCCTGCC CCCATCCCGG 1140

GATGAGCTGA CCAAGAACCA GGTCAGCCTG ACCTGCCTGG TCAAAGGCTT CTATCCCAGC 1200

GACATCGCCG TGGAGTGGGA GAGCAATGGG CAGCCGGAGA ACAACTACAA GACCACGCCT 1260

C[illegible]GTGCTGG ACTCCGACGG CTCCTTCTTC CTCTACAGCA AGCTCACCGT GGACAAGAGC 1320

AGGTGGCAGC AGGGGAACGT CTTCTCATGC TCCGTGATGC ATGAGGCTCT GCACAACCAC 1380

TACACGCAGA AGAGCCTCTC CCTGTCTCCG GGTAAATGAT AGATATC 1427

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ser Asn Phe Ile Gln Tyr Ser

1 5

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 762 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
GAATTCCATG GACATGAGGG TCCCCGCTCA GCTCCTAGGG CTCCTGCTGC TCTGGCTCCG    60
AGGTGCCAGA TGTGACATCC AGATGACCCA GTCTCCATCC TCCCTGTCTG CATCTGTAGG   120
AGACAGAGTC ACCATCACTT GCCGGGCAAC TCAGAGTGTT AGTAACTTTT TAAATTGGTA   180
TCAGCAGAAG CCAGGGGAAG CCCCTACGCT CCTGATCTAT GATGCATCCA CTTCGCAAAG   240
TGGGGTCCCA TCAAGGTTCA GTGGCAGTGG ATCTGGGATG GATTTCAGTC TCACCATCAG   300
CAGTCTGCAG CCTGAAGATC TTGCAATGTA TTACTGTCAA GCGAGTATCA ATACCCCGCT   360
TTTCGGCGGA GGGACCAGAA TAGATATGAG ACGAACTGTG GCTGCACCAT CTGTCTTCAT   420
CTTCCCGCCA TCTGATGAGC AGTTGAAATC TGGAACTGCC TCTGTTGTGT GCCTGCTGAA   480
TAACTTCTAT CCCAGAGAGG CCAAAGTACA GTGGAAGGTG GATAACGCCC TCCAATCGGG   540
TAACTCCCAG GAGAGTGTCA CAGAGCAGGA CAGCAAGGAC AGCACCTACA GCCTCAGCAG   600
CACCCTGACG CTGAGCAAAG CAGACTACGA GAAACACAAA GTCTACGCCT GCGAAGTCAC   660
CCATCAGGGC CTGAGCTCGC CCGTCACAAA GAGCTTCAAC AGGGGAGAGT GTTAGTGAGA   720
TGATCCTCTA GATCTACGTA TGATCAGCCT CGACTGTGCC TT                       762
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Ser Pro Val Thr Lys Ser Phe Thr Arg Gly Gln Cys
 1               5                  10
```

What is claimed is:

1. A reshaped human monoclonal antibody and functional fragments thereof, specifically reactive with an F protein epitope of Respiratory Syncytial Virus and capable of neutralizing infection by said virus selected from the group consisting of Hu19A, Hu19B, Hu19C and Hu19D.

2. The monoclonal antibody or functional fragment thereof according to Claim 1 which comprises a light chain amino acid sequence of Figure 3 selected from Sequences 19A, 19B, 19C and 19D and/or a heavy chain amino acid sequence of Figure 2 selected from Sequences 19A, 19B, 19C and 19D.

3. The monoclonal antibody according to Claim 1 wherein said fragment is selected from the group consisting of Fv, Fab and $F(ab')_2$.

4. An isolated nucleic acid molecule selected from the group consisting of:

(a) a nucleic acid sequence encoding any of the human monoclonal antibodies and functional fragments thereof of claim 1;

(b) a nucleic acid complementary to any of the sequences in (a); and

(c) a nucleic acid sequence of 18 or more nucleotides capable of hybridizing to (a) or (b) under stringent conditions.

5. A isolated nucleic acid molecule encoding a monoclonal antibody or functional fragment thereof according to Claim 1 having a nucleotide sequence of Figure 4.

6. A recombinant plasmid comprising a nucleic acid sequence of Claim 4.

7. A recombinant plasmid comprising a nucleic acid sequence of claim 5.

8. A plasmid according to Claim 7 encoding a protein sequence of Figure 2 or 3.

9. A host cell comprising the plasmid of Claim 8.

10. A process for the production of a human antibody specific for RSV comprising culturing the host cell of Claim 9 in a medium under suitable conditions of time, temperature and pH and recovering the antibody so produced.

11. A method of detecting RSV comprising contacting a source suspected of containing RSV with a diagnostically effective amount of the monoclonal antibody of Claim 1 and determining whether the monoclonal antibody binds to the source.

12. A method for providing passive immunotherapy to RSV disease in a human, comprising administering to the human an immunotherapeutically effective amount of the monoclonal antibody of Claim 1.

13. The method according to Claim 12 wherein the passive immunotherapy is provided prophylactically.

14. A pharmaceutical composition comprising at least one dose of an immunotherapeutically effective amount of the monoclonal antibody of Claim 1 in a pharmaceutically acceptable carrier.

15. A pharmaceutical composition comprising at least one dose of an immunotherapeutically effective amount of the monoclonal antibody of Claim 1 in combination with at least one additional monoclonal antibody.

16. The pharmaceutical composition according to Claim 15 wherein said additional monoclonal antibody is an anti-RSV antibody distinguished from the antibody of Claim 1 by virtue of being reactive with a different epitope of the RSV F protein antigen.

FIGURE 1

Conversion of Hu19 Fab to a Complete IgG1 mAb

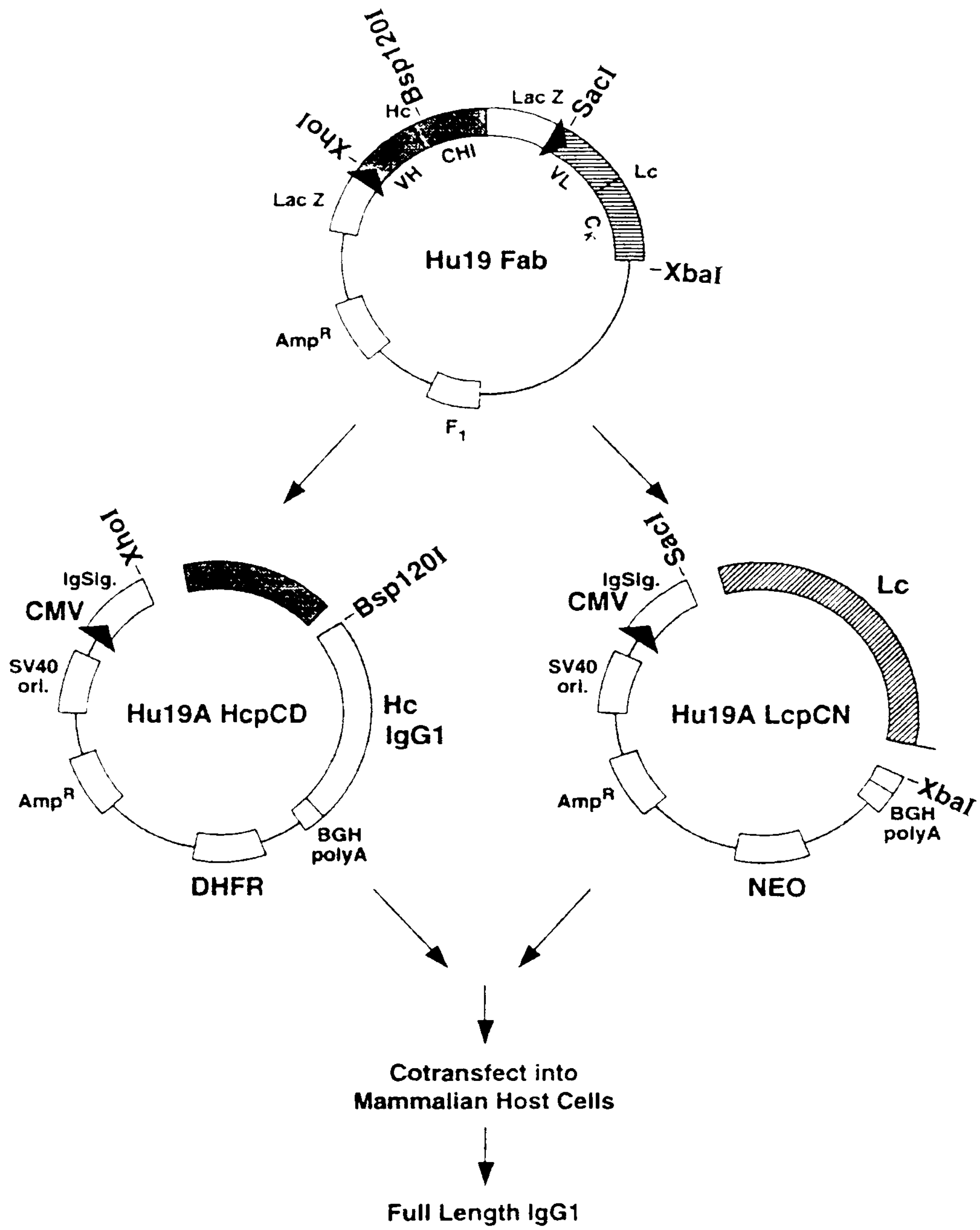

Figure 2

Comparison of the Heavy Chain Amino Acid Sequences of various Hu19 mAbs

```
                                             **
GL Dp58:                        EVQL VESGGGLVQPGGSLRLSCAASGFTFS     30
19A:        MGWSCIILFLVATATGVHS----LE-------R-------------T-L-
19B:        -EFGLSWV----LLR--QCQVQL V-------------------------
19C:        ---------------------- --------------------------
19D:        ---------------------- --------------------------


            CDR1                      CDR2
            -----               -----------------
GL Dp58     SYEMNWVRQAPGKGLEWVSYISSSGSTIYYADSVKGRFTISRDNAKNSLY     80
19A         G-T-H--------------S-TGGSNF-N-S------------------
19B         --------------------------------------------------
19C:        -------------------------------A------------------
19D:        -----------------------------Q-S----------------


                          CDR3
                       ----------
GL: Dp58    LQMNSLRAEDTAVYYCAR  98                      (SEQ ID NO: 4)
19A:        ------T----------TAPIAPPYFDHWGQGTLVTVSS     (SEQ ID NO: 5)
19B:        ---------------------------------------     (SEQ ID NO: 6)
19C:        ---------------------------------------     (SEQ ID NO: 7)
19D:        ---------------------------------------     (SEQ ID NO: 8)
```

Figure 3

Comparison of the Light Chain Amino Acid Sequences of various Hu19 mAbs

A. Leader and Variable

		CDR1	
GL Dpk9		DIQMTQSPSSLSASVGDRVTITC**RASQSIS**	30
19A	**MGWSCIILFLVATATGVHS**	EL-------------------**T--V-**	28
19B,C,D	**MRVPAQLLGLLLLWLRGARC**	DIQM--------------------------	

	CDR2	
GL Dpk9	**SYLN**WYQQKPGKAPKLLIY**AASSLQS**GVPSRFSGSGSGTDFTLTISSLQP	80
19A	**NFLN**------E--T----**D--TS**---------------M--S--------	78
19B,C,D	--	

	CDR3			
GL Dpk9	EDFATYYC	*		(SEQ ID NO: 9)
19A	--L-M---**QASINTPL**FGGGTRIDMRR		105	(SEQ ID NO: 10)
19B,C,D	----------------------------			(SEQ ID NO: 11)

B. Constant Region (Cκ)

Hu-κ,19C,D:	TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGN	
19A,B:	--	
Hu-κ,19C,D	SQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKS	
19A,B	---L----	
Hu-κ,19C,D	FNRGEC	(SEQ ID NO: 12)
19A,B	------	(SEQ ID NO: 13)

Figure 4

Figure 4A-- DNA sequence of the plasmid Hu19AHcpcd

```
  1  gacgtcgcggccgctctaggcctccaaaaaagcctcctcactacttctgg   50
 51  aatagctcagaggccgaggcggcctcggcctctgcataaataaaaaaat  100
101  tagtcagccatgcatggggcggagaatgggcggaactgggcggagttagg  150
151  ggcgggatgggcggagttaggggcgggactatggttgctgactaattgag  200
201  atgcatgctttgcatacttctgcctgctggggagcctggggactttccac  250
251  acctggttgctgactaattgagatgcatgctttgcatacttctgcctgct  300
301  ggggagcctggggactttccacaccctaactgacacacattccacagaat  350
351  taattcccggggatcgatccgtcgacgtacgactagttattaatagtaat  400
401  caattacggggtcattagttcatagcccatatatggagttccgcgttaca  450
451  taacttacggtaaatggcccgcctggctgaccgcccaacgacccccgccc  500
501  attgacgtcaataatgacgtatgttcccatagtaacgccaatagggactt  550
551  tccattgacgtcaatgggtggactatttacggtaaactgcccacttggca  600
601  gtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatga  650
651  cggtaaatggcccgcctggcattatgcccagtacatgaccttatgggact  700
701  ttcctacttggcagtacatctacgtattagtcatcgctattaccatggtg  750
751  atgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacg  800
```

[illegible]

[illegible]

901 [illegible] 950

Eco RI

951 tgggtacgtgaaccgtcagatcgcctggagacgccatcgaattctgagca 1000

1001 cacaggacctcaccatgggatggagctgtatcatcctcttcttggtagca 1050
M G W S C I I L F L V A
Leader start

1051 acagctacaggtgtccactccgaggtccaactgctcgaggagtctgggg 1100
T A T G V H S E V Q L L E V - (SEQ ID NO: 1[illegible])
Processed N-term.

1101 aggcctggtcaggcctggcgggtccctaagactctcgtgtgcagcctctg 1150

1151 gaaccaccctcagtggctataccatgcactgggtccgccaggctccaggg 1200

1201 aaggggctggagtgggtctcatccattactggaggtagcaacttcataaa 1250

1251 ctactcagactcagtgaagggccgattcaccatctccagagacaacgcca 1300

1301 agaactcactttatctgcaaatgaacagcctgacagccgaggacacggct 1350

1351 gtctattattgtgcgaccgcccctatagcaccgcccctactttgaccactg 1400

1401 gggccagggaaccctggtcaccgtctcctcagcctccaccaagggcccat 1450

1451 cggtcttccccctggcaccctcctccaagagcacctctgggggcacagcg 1500

1501 gccctgggctgcctggtcaaggactacttccccgaaccggtgaccgtgtc 1550

1551 gtggaactcaggcgccctgaccagcggcgtgcacaccttcccggctgtcc 1600

1601 tacagtcctcaggactctactccctcagcagcgtggtgactgtgccctcc 1650

1651 agcagcttgggcacccagacctacatctgcaacgtgaatcacaagcccag 1700

1701 caacaccaaggtggacaagaaagttgagcccaaatcttgtgacaaaactc 1750

1751 acacatgcccaccgtgcccagcacctgaactcctggggggaccgtcagtc 1800

1801 ttcctcttccccccaaaacccaaggacaccctcatgatctcccggacccc 1850

1851 tgaggtcacatgcgtggtggtggacgtgagccacgaagaccctgaggtca 1900

1901 agttcaactggtacgtggacggcgtggaggtgcataatgccaagacaaag 1950

1951 ccgcgggaggagcagtacaacagcacgtaccgggtggtcagcgtcctcac 2000

2001 cgtcctgcaccaggactggctgaatggcaaggagtacaagtgcaaggtct 2050

2051 ccaacaaagccctcccagcccccatcgagaaaaccatctccaaagccaaa 2100

2101 gggcagccccgagaaccacaggtgtacaccctgcccccatcccgggatga 2150

2151 gctgaccaagaaccaggtcagcctgacctgcctggtcaaaggcttctatc 2200

2201 ccagcgacatcgccgtggagtgggagagcaatgggcagccggagaacaac 2250

2251 tacaagaccacgcctcccgtgctggactccgacggctccttcttcctcta 2300

2301 cagcaagctcaccgtggacaagagcaggtggcagcaggggaacgtcttct 2350

[illegible]

[illegible]

[illegible] heavy chain [illegible]

[illegible]

2501 ttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatg 2550

2551 aggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggt 2600

2601 ggggtggggcaggacagcaagggggaggattgggaagacaatagcaggca 2650

2651 tgctggggatgcggtgggctctatggaaccagctggggctcgacagcgct 2700

2701 ggatctcccgatccccagctttgcttctcaatttcttatttgcataatga 2750

2751 gaaaaaaaggaaaattaattttaacaccaattcagtagttgattgagcaa 2800

2801 atgcgttgccaaaaaggatgctttagagacagtgttctctgcacagataa 2850

2851 ggacaaacattattcagagggagtacccagagctgagactcctaagccag 2900

2901 tgagtggcacagcattctagggagaaatatgcttgtcatcaccgaagcct 2950

2951 gattccgtagagccacaccttggtaagggccaatctgctcacacaggata 3000

3001 gagagggcaggagccagggcagagcatataaggtgaggtaggatcagttg 3050

3051 ctcctcacatttgcttctgacatagttgtgttgggagcttggatagcttg 3100

3101 gacagctcagggctgcgatttcgcgccaaacttgacggcaatcctagcgt 3150

3151 gaaggctggtaggattttatccccgctgccatcatggttcgaccattgaa 3200

3201 ctgcatcgtcgccgtgtcccaaaatatggggattggcaagaacggagacc 3250

3251 taccctggcctccgctcaggaacgagttcaagtacttccaaagaatgacc 3300

3301 acaacctcttcagtggaaggtaaacagaatctggtgattatgggtaggaa 3350

3351 aacctggttctccattcctgagaagaatcgacctttaaaggacagaatta 3400

3401 atatagttctcagtagagaactcaaagaaccaccacgaggagctcatttt 3450

3451 cttgccaaaagtttggatgatgccttaagacttattgaacaaccggaatt 3500

3501 ggcaagtaaagtagacatggtttggatagtcggaggcagttctgtttacc 3550

3551 aggaagccatgaatcaaccaggccaccttagactctttgtgacaaggatc 3600

3601 atgcaggaatttgaaagtgacacgtttttcccagaaattgatttggggaa 3650

3651 atataaacttctcccagaatacccaggcgtcctctctgaggtccaggagg 3700

3701 aaaaaggcatcaagtataagtttgaagtctacgagaagaaagactaacag 3750

3751 gaagatgctttcaagttctctgctcccctcctaaagctatgcatttttat 3800

3801 aagaccatgggacttttgctggctttagatcagcctcgactgtgccttct 3850

3851 agttgccagccatctgttgtttgcccctcccccgtgccttccttgaccct 3900

3901 ggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcat 3950

3951 cgcattgtctgagtaggtgtcattctattctggggggtggggtggggcag 4000

4001 gacagcaagggggaggattgggaagacaatagcaggcatgctggggatgc 4050

4051 [illegible] 4100

4101 [illegible] 4150

4151 gtgaaattgttatccgctcacaattccacacaacatacgagccggaagca 4200

4201 taaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaatt 4250

4251 gcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagct 4300

4301 gcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggc 4350

4351 gctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctg 4400

4401 cggcgagcggtatcagctcactcaaaggcggtaatacggttatccacaga 4450

4451 atcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaagg 4500

4501 ccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgc 4550

4551 ccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaa 4600

4601 cccgacaggactataaagataccaggcgtttccccctggaagctccctcg 4650

4651 tgcgctctcctgttccgaccctgccgcttaccggatacctgtccgccttt 4700

4701 ctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatct 4750

4751 cagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccc 4800

4801 ccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtcc 4850

4851 aacccggtaagacacgacttatcgccactggcagcagccactggtaacag 4900

4901 gattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggt 4950

4951 ggcctaactacggctacactagaaggacagtatttggtatctgcgctctg 5000

5001 ctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaa 5050

5051 acaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagatta 5100

5101 cgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacgggg 5150

5151 tctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgag 5200

5201 attatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtt 5250

5251 ttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaa 5300

5301 tgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatc 5350

5351 catagttgcctgactccccgtcgtgtagataactacgatacgggagggct 5400

5401 taccatctggccccagtgctgcaatgataccgcgagacccacgctcaccg 5450

5451 gctccagatttatcagcaataaaccagccagccggaagggccgagcgcag 5500

5501 aagtggtcctgcaactttatccgcctccatccagtctattaattgttgcc 5550

5551 gggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgtt 5600

5601 gccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttc 5650

5651 attcagctccggttcccaacgatcaaggcgagttacatgatcccccatgt 5700

5701 tgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagt 5750

[illegible]

[illegible]

[illegible]

5901 ccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagt 5950

5951 gctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttac 6000

6001 cgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatct 6050

6051 tcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaag 6100

6101 gcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatac 6150

6151 tcatactcttcctttttcaatattattgaagcatttatcagggttattgt 6200

6201 ctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagg 6250

6251 ggttccgcgcacatttccccgaaaagtgccacct 6284 (SEQ ID NO:14)

Figure 4B-- DNA sequence of the plasmid Hu19ALcpcn:

1	gacgtcgcggccgctctaggcctccaaaaaagcctcctcactacttctgg	50
51	aatagctcagaggccgaggcggcctcggcctctgcataaataaaaaaaat	100
101	tagtcagccatgcatggggcggagaatgggcggaactgggcggagttagg	150
151	ggcgggatgggcggagttaggggcgggactatggttgctgactaattgag	200
201	atgcatgctttgcatacttctgcctgctggggagcctggggactttccac	250
251	acctggttgctgactaattgagatgcatgctttgcatacttctgcctgct	300
301	ggggagcctggggactttccacaccctaactgacacacattccacagaat	350
351	taattcccggggatcgatccgtcgacgtacgactagttattaatagtaat	400
401	caattacggggtcattagttcatagcccatatatggagttccgcgttaca	450
451	taacttacggtaaatggcccgcctggctgaccgcccaacgacccccgccc	500
501	attgacgtcaataatgacgtatgttcccatagtaacgccaatagggactt	550
551	tccattgacgtcaatgggtggactatttacggtaaactgcccacttggca	600
601	gtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatga	650
651	cggtaaatggcccgcctggcattatgcccagtacatgaccttatgggact	700
701	ttcctacttggcagtacatctacgtattagtcatcgctattaccatggtg	750
751	atgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacg	800

801 [illegible] 850

851 [illegible] 900

901 [illegible] 950

Eco RI

951 ttggtacctgaaccatcagatcacctggaaacggtatcgaattcgagca 1000

1001 cacaggacctcaccatgggatggagctgtatcatcctcttcttggtagca 1050

M G W S C I I L F L V A

Leader start

1051 acagctacaggtgtccactccgagctcacccagtctccatcctccctgtc 1100

T A T G V H S E L T Q S P - (SEQ ID NO: 1[illegible])

Processed N-term.

1101 tgcatctgtaggagacagagtcaccatcacttgccgggcaactcagagtg 1150

1151 ttagtaactttttaaattggtatcagcagaagccaggggaagcccctacg 1200

1201 ctcctgatctatgatgcatccactttgcaaagtggggtcccatcaaggtt 1250

1251 cagtggcagtggatctgggatggatttcagtctcaccatcagcagtctgc 1300

1301 agcctgaagatcttgcaatgtattactgtcaacagagtatcaataccccg 1350

1351 cttttcggcggagggaccagaatagatattaagacgaactgtggctgcacc 1400

1401 atctgtcttcatcttcccgccatctgatgagcagttgaaatctggaactg 1450

1451 cctctgttgtgtgcctgctgaataacttctatcccagagaggccaaagta 1500

1501 cagtggaaggtggataacgccctccaatcgggtaactcccaggagagtgt 1550

```
1551  cacagagcaggacagcaaggacagcacctacagcctcagcagcaccctga  1600

1601  cgctgagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtc  1650

1651  acccatcagggcctgagcttgcccgtcacaaagagcttcaacaggggaga  1700
                                L  P  V  T  K  S  F  N  R  G  E

                              Xba I
1701  gtgttagtgagatgatcctctagagtcatctacgtatgatcagcctcgac  1750
        C  *  end of light chain       (SEQ ID NO: 19)

1751  tgtgccttctagttgccagccatctgttgtttgcccctcccccgtgcctt  1800

1801  ccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgag  1850

1851  gaaattgcatcgcattgtctgagtaggtgtcattctattctggggggtgg  1900

1901  ggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatg  1950

1951  ctggggatgcggtgggctctatggaaccagctggggctcgacagctcgag  2000

2001  ctagctttgcttctcaatttcttatttgcataatgagaaaaaaaggaaaa  2050

2051  ttaattttaacaccaattcagtagttgattgagcaaatgcgttgccaaaa  2100

2101  aggatgctttagagacagtgttctctgcacagataaggacaaacattatt  2150

2151  cagagggagtacccagagctgagactcctaagccagtgagtggcacagca  2200

2201  ttctagggagaaatatgcttgtcatcaccgaagcctgattccgtagagcc  2250

2251  acaccttggtaagggccaatctgctcacacaggatagagagggcaggagc  2300
```

2301 ctgggcagaccatataaggtgaggtaggatcagttgctcctcacatttct 2350

2351 ttctgacatagttatgttaggagcttagatcgatccaccatgattgaaca 2400

2401 agatggattgcacgcaggttctccggccgcttgggtggagaggctattcg 2450

2451 gctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttc 2500

2501 cggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtc 2550

2551 cggtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctgg 2600

2601 ccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcg 2650

2651 ggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtc 2700

2701 atctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgc 2750

2751 ggcggctgcatacgcttgatccggctacctgcccattcgaccaccaagcg 2800

2801 aaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcga 2850

2851 tcaggatgatctggacgaagagcatcaggggctcgcgccagccgaactgt 2900

2901 tcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacc 2950

2951 catggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttc 3000

3001 tggattcatcgactgtggccggctgggtgtggcggaccgctatcaggaca 3050

3051 tagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggct 3100

3101 gaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcat 3150

3151 cgccttctatcgccttcttgacgagttcttctgagcgggactctggggtt 3200

3201 cgaaatgaccgaccaagcgacgcccaacctgccatcacgagatttcgatt 3250

3251 ccaccgccgccttctatgaaaggttgggcttcggaatcgttttccgggac 3300

3301 gccggctggatgatcctccagcgcggggatctcatgctggagttcttcgc 3350

3351 ccaccccaacttgtttattgcagcttataatggttacaaataaagcaata 3400

3401 gcatcacaaatttcacaaataaagcatttttttcactgcattctagttgt 3450

3451 ggtttgtccaaactcatcaatgtatcttatcatgtctggatcgcggccgc 3500

3501 gatcccgtcgagagcttggcgtaatcatggtcatagctgtttcctgtgtg 3550

3551 aaattgttatccgctcacaattccacacaacatacgagccggaagcataa 3600

3601 agtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcg 3650

3651 ttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgca 3700

3701 ttaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgct 3750

3751 cttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcgg 3800

3801 cgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatc 3850

3851 aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca 3900

3901 ggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccc 3950

3951 cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaaccc 4000

4001 gacaggactataaagataccaggcgtttccccctggaagctccctcgtgc 4050

4051 gctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctc 4100

4101 ccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcag 4150

4151 ttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccg 4200

4201 ttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaac 4250

4251 ccggtaagacacgacttatcgccactggcagcagccactggtaacaggat 4300

4301 tagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggc 4350

4351 ctaactacggctacactagaaggacagtatttggtatctgcgctctgctg 4400

4401 aagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca 4450

4451 aaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgc 4500

4501 gcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtct 4550

4551 gacgctcagtggaacgaaaactcacgttaagggattttggtcatgagatt 4600

4601 atcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtttta 4650

4651 aatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgc 4700

4701 ttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccat 4750

4751 agttgcctgactccccgtcgtgtagataactacgatacgggagggcttac 4800

4801 catctggccccagtgctgcaatgataccgcgagacccacgctcaccggct 4850

4851 ccagatttatcagcaataaaccagccagccggaagggccgagcgcagaag 4900

4901 tggtcctgcaactttatccgcctccatccagtctattaattgttgccggg 4950

4951 aagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgcc 5000

5001 attgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcatt 5050

5051 cagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgt 5100

5101 gcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaag 5150

5151 ttggccgcagtgttatcactcatggttatggcagcactgcataattctct 5200

5201 tactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaa 5250

5251 ccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccg 5300

5301 gcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgct 5350

5351 catcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgc 5400

5401 tgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttca 5450

5451 gcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggca 5500

5501 aaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactca 5550

5551 tactcttcctttttcaatattattgaagcatttatcagggttattgtctc 5600

5601 atgagcggatacatatttgaatgtatttagaaaaataaacaaataggggt 5650

5651 tccgcgcacatttccccgaaaagtgccacct 5681 (SEQ ID NO: 17)

Figure 4C Heavy chain coding sequence in the plasmid Hu19BHeped

Eco RI

gaattc**ggtacc** 1000

1001 **atggagtttgggctgagctgggttttcctcgtggctcttttaagaggtgt** 1050

M E F G L S W V F L V A L L R G V

Leader start

1051 **ccagtgtcaggtgcagctggt**ggagtctgggggaggcctggtcaggcctg 1100

Q C Q V Q L V - (SEQ ID NO: 21)

Processed N-term

1101 gcgggtccctaagactctcgtgtgcagcctctggaaccaccctcagtggc 1150

1151 tataccatgcactgggtccgccaggctccagggaaggggctggagtgggt 1200

1201 ctcatccattactggaggtagcaacttcataaactactcagactcagtga 1250

1251 agggccgattcaccatctccagagacaacgccaagaactcactttatctg 1300

1301 caaatgaacagcctgacagccgaggacacggctgtctattattgtgcgac 1350

1351 cgcccctatagcaccgccctactttgaccactggggccagggaaccctgg 1400

1401 tcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggca 1450

1451 ccctcctccaagagcacctctgggggcacagcggccctgggctgcctggt 1500

1501 caaggactacttccccgaaccggtgaccgtgtcgtggaactcaggcgccc 1550

1551 tgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactc 1600

1601 tactccctcagcagcgtggtgaccgtgccctccagcagcttgggcaccca 1650

1651 gacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggaca 1700

1701 agaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgc 1750

1751 ccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaa 1800

1801 acccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtgg 1850

1851 tggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtg 1900

1901 gacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagta 1950

1951 caacagcacgtaccgggtggtcagcgtcctcaccgtcctgcaccaggact 2000

2001 ggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctccca 2050

2051 gcccccatcgagaaaaccatctccaaagccaaagggcagccccgagaacc 2100

2101 acaggtgtacaccctgcccccatcccgggatgagctgaccaagaaccagg 2150

2151 tcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtg 2200

2201 gagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcc 2250

2251 cgtgctggactccgacggctccttcttcctctacagcaagctcaccgtgg 2300

2301 acaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcat 2350

2351 gaggctctgcacaaccactacacgcagaagagcctctccctgtctccggg 2400

S P G

2401 taaa<u>tga</u>tagatatc - (SEQ ID NO:20)

K * end of heavy chain

Figure 4D Light chain coding sequence in the plasmid Hu19BLcpen

Eco RI

gaattc**catgga** 1000

1001 **catgagggtccccgctcagctcctagggctcctgctgctctggctccgag** 1050

M R V P A Q L L G L L L L W L R

Leader start

1051 **gtgccagatgtgacatccagatg**acccagtctccatcctccctgtctgca 1100

G A R C D I Q M T - (SEQ ID NO: 23)

Processed N-term

1101 tctgtaggagacagagtcaccatcacttgccgggcaactcagagtgttag 1150

1151 taactttttaaattggtatcagcagaagccaggggaagcccctacgctcc 1200

1201 tgatctatgatgcatccacttcgcaaagtggggtcccatcaaggttcagt 1250

1251 ggcagtggatctgggatggatttcagtctcaccatcagcagtctgcagcc 1300

1301 tgaagatcttgcaatgtattactgtcaagcgagtatcaataccccgcttt 1350

1351 tcggcggagggaccagaatagatatgagacgaactgtggctgcaccatct 1400

1401 gtcttcatcttcccgccatctgatgagcagttgaaatctggaactgcctc 1450

1451 tgttgtgtgcctgctgaataacttctatcccagagaggccaaagtacagt 1500

1501 ggaaggtggataacgccctccaatcgggtaactcccaggagagtgtcaca 1550

1551 gagcaggacagcaaggacagcacctacagcctcagcagcaccctgacgct 1600

1601 gagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtcaccc 1650

1651 atcagggcctgagcttgcccgtcacaaagagcttcaacaggggagagtgt 1700

L P V T K S F N R G E C (SEQ NO: 24)

Xba I

1701 tagtgagatgatcctctaga (SEQ ID NO: 22)

* end of light chain

Figure 4E Heavy chain coding sequence in the plasmid Hu19CHeped

Eco RI

gaattc**ggtacc** 1000

1001 **atggagtttgggctgagctgggttttcctcgtggctcttttaagaggtgt** 1050

M E F G L S W V F L V A L L R G V

1051 **ccagtgtcaggtgcagctggt**ggagtctgggggaggcctggtcaggcctg 1100

Q C Q V Q L V - (SEQ ID NO: 21)

Processed N-term

1101 gcgggtccctaagactctcgtgtgcagcctctggaaccaccctcagtggc 1150

1151 tataccatgcactgggtccgccaggctccagggaaggggctggagtgggt 1200

1201 ctcatccattactggaggtagcaacttcataaactac**g**cagactcagtga 1250

S N F I N Y A - (SEQ ID NO: 26)

1251 agggccgattcaccatctccagagacaacgccaagaactcactttatctg 1300

1301 caaatgaacagcctgacagccgaggacacggctgtctattattgtgcgac 1350

1351 cgcccctatagcaccgccctactttgaccactggggccagggaaccctgg 1400

1401 tcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggca 1450

1451 ccctcctccaagagcacctctgggggcacagcggccctgggctgcctggt 1500

1501 caaggactacttccccgaaccggtgacggtgtcgtggaactcaggcgccc 1550

1551 tgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactc 1600

1601 tactccctcagcagcgtggtgaccgtgccctccagcagcttgggcaccca 1650

1651 gacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggaca 1700

1701 agaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgc 1750

1751 ccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaa 1800

1801 acccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtgg 1850

1851 tggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtg 1900

1901 gacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagta 1950

1951 caacagcacgtaccgggtggtcagcgtcctcaccgtcctgcaccaggact 2000

2001 ggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctccca 2050

2051 gcccccatcgagaaaaccatctccaaagccaaagggcagccccgagaacc 2100

2101 acaggtgtacaccctgcccccatcccgggatgagctgaccaagaaccagg 2150

2151 tcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtg 2200

2201 gagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcc 2250

2251 cgtgctggactccgacggctccttcttcctctacagcaagctcaccgtgg 2300

2301 acaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcat 2350

2351 gaggctctgcacaaccactacacgcagaagagcctctccctgtctccggg 2400

S P G

2401 taaatgatagatatc - (SEQ ID NO: 25)

K * end of heavy chain

Figure 4F Heavy chain coding sequence in the plasmid Hu19DHcpcd

Eco RI

gaattc**ggtacc** 1000

1001 **atggagtttgggctgagctgggttttcctcgtggctcttttaagaggtgt** 1050

M E F G L S W V F L V A L L R G V

1051 **ccagtgtcaggtgcagctggt**ggagtctggggaggcctggtcaggcctg 1100

Q C Q V Q L V - (SEQ ID NO: 21)

Processed N-term

1101 gcgggtccctaagactctcgtgtgcagcctctggaaccaccctcagtggc 1150

1151 tataccatgcactgggtccgccaggctccagggaaggggctggagtgggt 1200

1201 ctcatccattactggaggtagcaacttcata**c**a**a**tactcagactcagtga 1250

S N F I Q Y S - (SEQ ID NO: 28)

1251 agggccgattcaccatctccagagacaacgccaagaactcactttatctg 1300

1301 caaatgaacagcctgacagccgaggacacggctgtctattattgtgcgac 1350

1351 cgcccctatagcaccgccctactttgaccactggggccagggaaccctgg 1400

1401 tcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggca 1450

1451 ccctcctccaagagcacctctgggggcacagcggccctgggctgcctggt 1500

1501 caaggactacttccccgaaccggtgaccgtgtcgtggaactcaggcgccc 1550

1551 tgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactc 1600

1601 tactccctcagcagcgtggtgaccgtgccctccagcagcttgggcaccca 1650

1651 gacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggaca 1700

1701 agaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgc 1750

1751 ccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaa 1800

1801 acccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtgg 1850

1851 tggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtg 1900

1901 gacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagta 1950

1951 caacagcacgtaccgggtggtcagcgtcctcaccgtcctgcaccaggact 2000

2001 ggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctccca 2050

2051 gcccccatcgagaaaaccatctccaaagccaaagggcagccccgagaacc 2100

2101 acaggtgtacaccctgcccccatcccgggatgagctgaccaagaaccagg 2150

2151 tcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtg 2200

2201 gagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcc 2250

2251 cgtgctggactccgacggctccttcttcctctacagcaagctcaccgtgg 2300

2301 acaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcat 2350

2351 gaggctctgcacaaccactacacgcagaagagcctctccctgtctccggg 2400

S P G

2401 taaatgatagatatc - (SEQ ID NO: 27)

K * end of heavy chain

Figure 4G Light chain coding sequence in the plasmid H19CLepen

Eco RI

gaattc**catgga** 1000

1001 **catgagggtccccgctcagctcctagggctcctgctgctctggctccgag** 1050

M R V P A Q L L G L L L L W L R

Leader start

1051 **gtgccagatgtgacatccagatg**acccagtctccatcctccctgtctgca 1100

G A R C D I Q M T - (SEQ ID NO: 23)

Processed N-term

1101 tctgtaggagacagagtcaccatcacttgccgggcaactcagagtgttag 1150

1151 taactttttaaattggtatcagcagaagccaggggaagcccctacgctcc 1200

1201 tgatctatgatgcatccacttcgcaaagtggggtcccatcaaggttcagt 1250

1251 ggcagtggatctgggatggatttcagtctcaccatcagcagtctgcagcc 1300

1301 tgaagatcttgcaatgtattactgtcaagcgagtatcaataccccgcttt 1350

1351 tcggcggagggaccagaatagatatgagacgaactgtggctgcaccatct 1400

1401 gtcttcatcttcccgccatctgatgagcagttgaaatctggaactgcctc 1450

1451 tgttgtgtgcctgctgaataacttctatcccagagaggccaaagtacagt 1500

1501 ggaaggtggataacgccctccaatcgggtaactcccaggagagtgtcaca 1550

1551 gagcaggacagcaaggacagcacctacagcctcagcagcaccctgacgct 1600

```
1601  gagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtcaccc  1650

1651  atcagggcctgagctcgcccgtcacaaagagcttcaacaggggagagtgt  1700
                      S   P   V   T   K   S   F   T   R   G   Q   C  (SEQ NO: 30)

                    Xba I
1701  tagtgagatgatcctctagatctacgtatgatcagcctcgactgtgcctt -(SEQ NO: 29)
       *  end of light chain
```

Fig. 5. COOMASSIE STAINED SDS-PAGE GEL ANALYSIS OF 10UG HU19B AND HU19C RESPECTIVELY UNDER REDUCING CONDITIONS

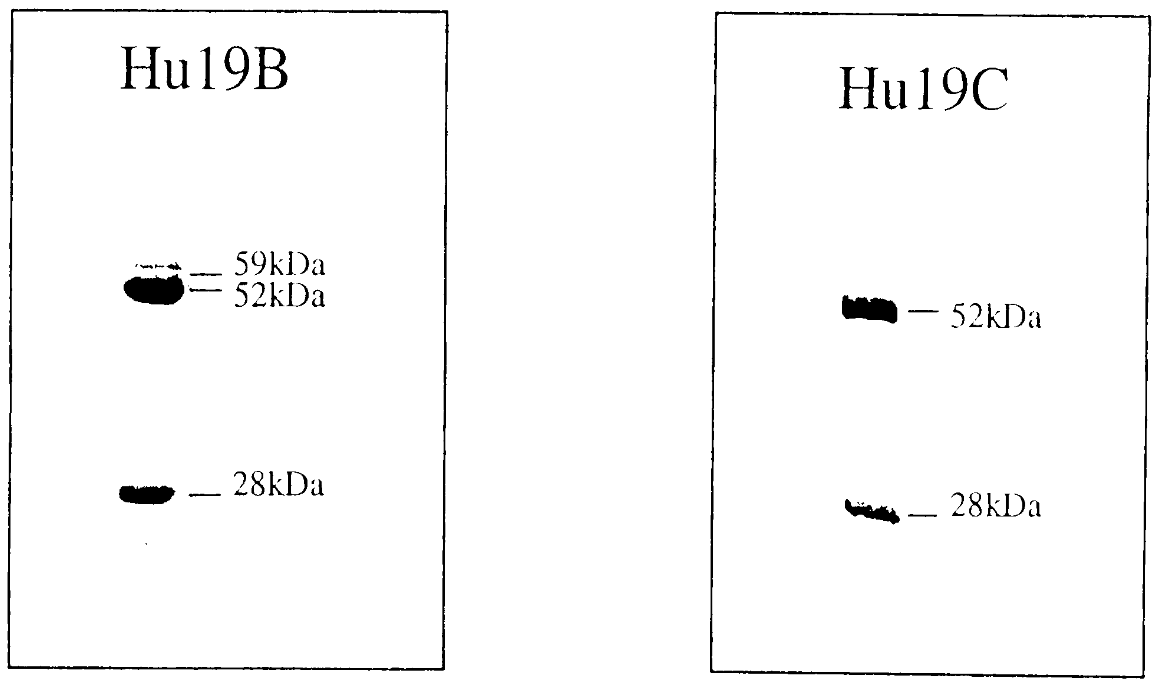

Fig. 6. SEPARATION OF HU19B GLYCOVARIANTS BY ANION EXCHANGE CHROMATOGRAPHY

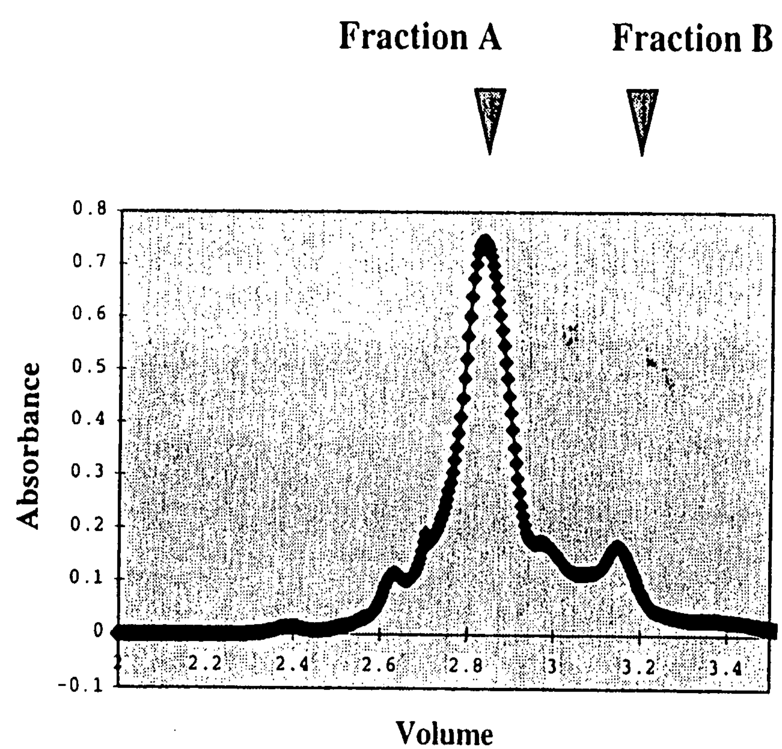

Fig.7. SDS-PAGE of Glycosylation Variant of Hu19B Fab

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/19203

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 39/395
US CL :530/387.3, 388.3; 424/159.1, 133.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/387.3, 388.3; 424/159.1, 133.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS. MEDLINE, DIALOG
search terms: respiratory syncytial virus, reshaped human monoclonal antibody

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TEMPTEST et al Reshaping a human monoclonal antibody to inhibit human respiratory syncytial virus infection in vivo. Bio/Technology. March 1994. Vol 9. Pages 266-271. See entire document.	1-11 and 14-16
A	GROOTHIUS et al. Prophylactic administration of respiratory syncytial virus immune globulin to high-risk infants and young childern. The New England Journal of Medicine. 18 November 1993. Volume 329, Number 21, pages 1524-1530. See entire document.	1-11 and 14-16

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 JANUARY 1998	23 FEB 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer JULIE E. REEVES
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/19203

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CROWE et al. Recombinant human respiratory syncytial virus (RSV) monoclonal antibody Fab is effective therapeuticaly when introduced directly into the lungs of RSV-infected mice. Proc. Natl. Acad. Sci. USA. February 1994. Vol 91. pages 1368-1390. See entire document.	1-11 and 14-16
A	WALSH et al. Protection from respiratory syncytial virus infection in cotton rats by passive transfer of monoclonal antibodies. Infection and Immunity. February 1984. Vol. 43. No. 2. pages 756-758. See entire document.	1-11 and 14-161-11
A	WALSH et al. Analysis of the respiratory syncytial virus fusion protein using monoclonal and polyclonal antibodies. J. Gen. Virol. 1986. Vol 67. pages 505-513. See entire document.	1-11 and 14-16
A	BARBAS III et al. Human monoclonal Fab fragments derived from a combinatorial library bind to respiratory syncytial virus F glycoprotein and neutralize infectivity. Proc. Natl. Acad. Sci. USA. November 1992. Vol 89. 10164-10168. See entire document.	1-11 and 14-161-11 a
A	SIBER et al. Comparison of antibody concentrations and protective activity of respiratory syncytial virus immune globulin and conventional immune globulin. Journal of Infectious Diseases. 1994. Vol 169. 1368-1373. See entire document.	1-11 and 14-16

FIGURE 1

Conversion of Hu19 Fab to a Complete IgG1 mAb

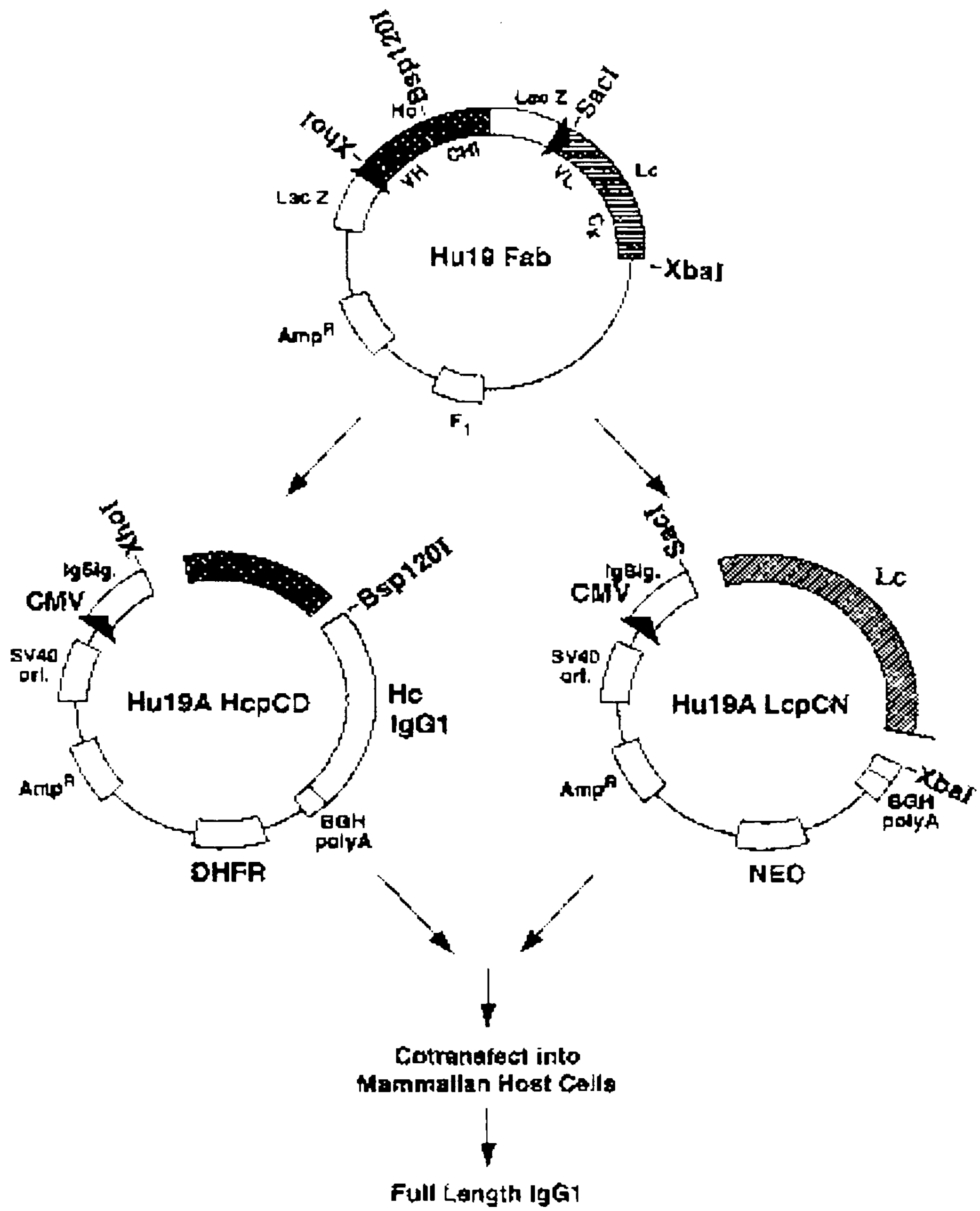

Figure 2

Comparison of the Heavy Chain Amino Acid Sequences of various Hu19 mAbs

```
                                                 **
GL Dp58:                       EVQL VESGGGLVQPGGSLRLSCAASGFTFS    30
19A:     MGWSCIILFLVATATGVHS-----LE    ----R------------T-L-
19B:     -EFGLSWV----LLR--QCQVQL V---------   . .      -------
19C:     ---------------------- --- .    .  . -----------------
19D:     ---------------------- ----------------------    ---

         CDR1                CDR2
         -----          ------------ --
GL Dp58  SYEMNWVRQAPGKGLEWVSYISSSGSTIYYADSVKGRFTISRDNAKNSLY    80
19A      G-T-H--------------S--TGGSNF-N-S------------------
19B      ------------------------   . ----------------------
19C      -----------------------------A-----------------   ---
19D      --- ..       ----------------Q-S-----------------

                    CDR3
                 ----------
GL: Dp58 LQMNSLRAEDTAVYYCAR  98                (SEQ ID NO: 4)
19A:     ------T----------TAPIAPPYFDHWGQGTLVTVSS  (SEQ ID NO: 5)
19B:     ---------------------------------------  (SEQ ID NO: 6)
19C:     ---------------------------------------  (SEQ ID NO: 7)
19D:     -- ------------------------------------  (SEQ ID NO: 8)
```

Figure 3

Comparison of the Light Chain Amino Acid Sequences of various Hu19 mAbs

A. Leader and Variable

```
                                                               CDR1
                                                              ------
GL Dpk9:                          DIQMTQSPSSLSASVGDRVTITCRASQSIS     30
19A:       MGWSCIILFLVATATGVHS    EL---------------------T--V-       26
19B,C,D:   MRVPAQLLGLLLLWLRGARCDIQM-------------------------

                                  CDR2
           ---               -------
GL Dpk9:   SYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTISSLQP    80
19A:       NFLN----    E  T ----D--TS---------------M---S--------    78
19B,C,D:                                   -----------    ---------

                    CDR3
                  --------
GL Dpk9:   EDFATYYC          *                                   (SEQ ID NO: 9)
19A:       --L-M---QASINTPLFGGGTRIDMRR        105                (SEQ ID NO: 10)
19B,C,D:   --------------------------                            (SEQ ID NO: 11)
```

B. Constant Region (Cκ)

```
Hu-κ,19C,D:  TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGN
19A,B:       --------------------------------------------------

Hu-κ,19C,D   SQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKS
19A,B:          ---------------------------       ----------L

Hu-κ,19C,D   FNRGEC          (SEQ ID NO: 12)
19A,B        ------          (SEQ ID NO: 13)
```

Figure 4

Figure 4A-- DNA sequence of the plasmid Hu19AHcpcd

```
  1  gacgtcggccgctctaggcctccaaaaaagcctcctcactacttctgg   50
 51  aatagctcagaggccgaggcggcctcggcctctgcataaataaaaaaat  100
101  tagtcagccatgcatggggcggagaatgggcggaactgggcggagttagg  150
151  ggcgggatgggcggagttaggggcgggactatggttgctgactaattgag  200
201  atgcatgctttgcatacttctgcctgctggggagcctggggactttccac  250
251  acctggttgctgactaattgagatgcatgctttgcatacttctgcctgct  300
301  ggggagcctggggactttccacaccctaactgacacacattccacagaat  350
351  taattcccggggatcgatccgtcgacgtacgactagttattaatagtaat  400
401  caattacggggtcattagttcatagcccatatatggagttccgcgttaca  450
451  taacttacggtaaatggcccgcctggctgaccgcccaacgacccccgccc  500
501  attgacgtcaataatgacgtatgttcccatagtaacgccaatagggactt  550
551  tccattgacgtcaatgggtggactatttacggtaaactgcccacttggca  600
601  gtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatga  650
651  cggtaaatggcccgcctggcattatgcccagtacatgaccttatgggact  700
701  ttcctacttggcagtacatctacgtattagtcatcgctattaccatggtg  750
751  atgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacg  800
```

801 gggatttccaagtctccaccccattgacgtcaatgggagtttgttttggc 850

851 accaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattg 900

901 acgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagc 950

Eco R1

951 tgggtacgtgaaccgtcagatcgcctggagacgccatcgaattctgagca 1000

1001 cacaggacctcaccatgggatggagctgtatcatcctcttcttggtagca 1050

M G W S C I I L F L V A

Leader start

1051 acagctacaggtgtccactccgaggtccaactgctcgaggagtctggggg 1100

T A T G V H S E V Q L L E V - (SEQ ID NO: 15)

Processed N-term.

1101 aggcctggtcaggcctggcgggtccctaagactctcgtgtgcagcctctg 1150

1151 gaaccaccctcagtggctataccatgcactgggtccgccaggctccaggg 1200

1201 aagggcctggagtgggtctcatccattactggaggtagcaacttcataaa 1250

1251 ctactcagactcagtgaagggccgattcaccatctccagagacaacgcca 1300

1301 agaactcactttatctgcaaatgaacagcctgacagccgaggacacggct 1350

1351 gtctattattgtgcgaccgcccctatagcaccgccctactttgaccactg 1400

1401 gggccagggaaccctggtcaccgtctcctcagcctccaccaagggcccat 1450

1451 cggtcttccccctggcaccctcctccaagagcacctctgggggcacagcg 1500

1501 gccctgggctgcctggtcaaggactacttccccgaaccggtgacggtgtc 1550

1551 gtggaactcaggcgccctgaccagcggcgtgcacaccttcccggctgtcc 1600

1601 tacagtcctcaggactctactccctcagcagcgtggtgactgtgccctcc 1650

1651 agcagcttgggcacccagacctacatctgcaacgtgaatcacaagcccag 1700

1701 caacaccaaggtggacaagaaagttgagcccaaatcttgtgacaaaactc 1750

1751 acacatgcccaccgtgcccagcacctgaactcctggggggaccgtcagtc 1800

1801 ttcctcttccccccaaaacccaaggacaccctcatgatctcccggacccc 1850

1851 tgaggtcacatgcgtggtggtggacgtgagccacgaagaccctgaggtca 1900

1901 agttcaactggtacgtggacggcgtggaggtgcataatgccaagacaaag 1950

1951 ccgcgggaggagcagtacaacagcacgtaccgggtggtcagcgtcctcac 2000

2001 cgtcctgcaccaggactggctgaatggcaaggagtacaagtgcaaggtct 2050

2051 ccaacaaagccctcccagcccccatcgagaaaaccatctccaaagccaaa 2100

2101 gggcagccccgagaaccacaggtgtacaccctgcccccatcccgggatga 2150

2151 gctgaccaagaaccaggtcagcctgacctgcctggtcaaaggcttctatc 2200

2201 ccagcgacatcgccgtggagtgggagagcaatgggcagccggagaacaac 2250

2251 tacaagaccacgcctcccgtgctggactccgacggctccttcttcctcta 2300

2301 cagcaagctcaccgtggacaagagcaggtggcagcaggggaacgtcttct 2350

2351 catgctccgtgatgcatgaggctctgcacaaccactacacgcagaagagc 2400

2401 ctctccctgtctccgggtaaatgatcgatatctacgtatgatcagcctcg 2450

S P G K * end of heavy chain (SEQ ID NO: 16)

2451 actgtgccttctagttgccagccatctgttgtttgcccctcccccgtgcc 2500

2501 ttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatg 2550

2551 aggaaattgcatcgcattgtctgagtaggtgtcattctattctggggggt 2600

2601 ggggtggggcaggacagcaagggggaggattgggaagacaatagcaggca 2650

2651 tgctggggatgcggtgggctctatggaaccagctggggctcgacagcgct 2700

2701 ggatctcccgatccccagctttgcttctcaatttcttatttgcataatga 2750

2751 gaaaaaaaggaaaattaattttaacaccaattcagtagttgattgagcaa 2800

2801 atgcgttgccaaaaaggatgctttagagacagtgttctctgcacagataa 2850

2851 ggacaaacattattcagagggagtacccagagctgagactcctaagccag 2900

2901 tgagtggcacagcattctagggagaaatatgcttgtcatcaccgaagcct 2950

2951 gattccgtagagccacaccttggtaagggccaatctgctcacacaggata 3000

3001 gagagggcaggagccagggcagagcatataaggtgaggtaggatcagttg 3050

3051 ctcctcacatttgcttctgacatagttgtgttgggagcttggatagcttg 3100

3101 gacagctcagggctgcgatttcgcgccaaacttgacggcaatcctagcgt 3150

3151 gaaggctggtaggattttatccccgctgccatcatggttcgaccattgaa 3200

3201 ctgcatcgtcgccgtgtcccaaaatatggggattggcaagaacggagacc 3250

3251 taccctggcctccgctcaggaacgagttcaagtacttccaaagaatgacc 3300

3301 acaacctcttcagtggaaggtaaacagaatctggtgattatgggtaggaa 3350

3351 aacctggttctccattcctgagaagaatcgacctttaaaggacagaatta 3400

3401 atatagttctcagtagagaactcaaagaaccaccacgaggagctcatttt 3450

3451 cttgccaaaagtttggatgatgccttaagacttattgaacaaccggaatt 3500

3501 ggcaagtaaagtagacatggtttggatagtcggaggcagttctgtttacc 3550

3551 aggaagccatgaatcaaccaggccacctcagactctttgtgacaaggatc 3600

3601 atgcaggaatttgaaagtgacacgtttttcccagaaattgatttggggaa 3650

3651 atataaacttctcccagaatacccaggcgtcctctctgaggtccaggagg 3700

3701 aaaaaggcatcaagtataagtttgaagtctacgagaagaaagactaacag 3750

3751 gaagatgctttcaagttctctgctcccctcctaaagctatgcatttttat 3800

3801 aagaccatgggacttttgctggctttagatcagcctcgactgtgccttct 3850

3851 agttgccagccatctgttgtttgcccctcccccgtgccttccttgaccct 3900

3901 ggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcat 3950

3951 cgcattgtctgagtaggtgtcattctattctggggggtggggtggggcag 4000

4001 gacagcaagggggaggattgggaagacaatagcaggcatgctggggatgc 4050

4051 ggtgggctctatggaaccagctggggctcgatcgagtgtatgactgcggc 4100

4101 cgcgatcccgtcgagagcttggcgtaatcatggtcatagctgtttcctgt 4150

4151 gtgaaattgttatccgctcacaattccacacaacatacgagccggaagca 4200

4201 taaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaatt 4250

4251 gcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagct 4300

4301 gcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggc 4350

4351 gctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctg 4400

4401 cggcgagcggtatcagctcactcaaaggcggtaatacggttatccacaga 4450

4451 atcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaagg 4500

4501 ccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgc 4550

4551 ccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaa 4600

4601 cccgacaggactataaagataccaggcgtttccccctggaagctccctcg 4650

4651 tgcgctctcctgttccgaccctgccgcttaccggatacctgtccgccttt 4700

4701 ctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatct 4750

4751 cagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccc 4800

4801 ccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtcc 4850

4851 aacccggtaagacacgacttatcgccactggcagcagccactggtaacag 4900

4901 gattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggt 4950

4951 ggcctaactacggctacactagaaggacagtatttggtatctgcgctctg 5000

5001 ctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaa 5050

5051 acaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagatta 5100

5101 cgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacgggg 5150

5151 tctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgag 5200

5201 attatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtt 5250

5251 ttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaa 5300

5301 tgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatc 5350

5351 catagttgcctgactccccgtcgtgtagataactacgatacgggagggct 5400

5401 taccatctggccccagtgctgcaatgataccgcgagacccacgctcaccg 5450

5451 gctccagatttatcagcaataaaccagccagccggaagggccgagcgcag 5500

5501 aagtggtcctgcaactttatccgcctccatccagtctattaattgttgcc 5550

5551 gggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgtt 5600

5601 gccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttc 5650

5651 attcagctccggttcccaacgatcaaggcgagttacatgatcccccatgt 5700

5701 tgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagt 5750

5751 aagttggccgcagtgttatcactcatggttatggcagcactgcataattc 5800

5801 tcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtact 5850

5851 caaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgc 5900

5901 ccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagt 5950

5951 gctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttac 6000

6001 cgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatct 6050

6051 tcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaag 6100

6101 gcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatac 6150

6151 tcatactcttcctttttcaatattattgaagcatttatcagggttattgt 6200

6201 ctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagg 6250

6251 ggttccgcgcacatttccccgaaaagtgccacct 6284 (SEQ ID NO:14)

Figure 4B-- DNA sequence of the plasmid Hu19ALcpcn:

1 gacgtcgcggccgctctaggcctccaaaaaagcctcctcactacttctgg 50

51 aatagctcagaggccgaggcggcctcggcctctgcataaataaaaaaaat 100

101 tagtcagccatgcatggggcggagaatgggcggaactgggcggagttagg 150

151 ggcgggatgggcggagttaggggcgggactatggttgctgactaattgag 200

201 atgcatgctttgcatacttctgcctgctggggagcctggggactttccac 250

251 acctggttgctgactaattgagatgcatgctttgcatacttctgcctgct 300

301 ggggagcctggggactttccacaccctaactgacacacattccacagaat 350

351 taattcccggggatcgatccgtcgacgtacgactagttattaatagtaat 400

401 caattacggggtcattagttcatagcccatatatggagttccgcgttaca 450

451 taacttacggtaaatggcccgcctggctgaccgcccaacgacccccgccc 500

501 attgacgtcaataatgacgtatgttcccatagtaacgccaatagggactt 550

551 tccattgacgtcaatgggtggactatttacggtaaactgcccacttggca 600

601 gtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatga 650

651 cggtaaatggcccgcctggcattatgcccagtacatgaccttatgggact 700

701 ttcctacttggcagtacatctacgtattagtcatcgctattaccatggtg 750

751 atgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacg 800

```
801   gggatttccaagtctccaccccattgacgtcaatgggagtttgttttggc   850

851   accaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattg   900

901   acgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagc   950

                                              Eco R1
951   tgggtacgtgaaccgtcagatcgcctggagacgccatcgaattctgagca   1000

1001  cacaggacctcaccatgggatggagctgtatcatcctcttcttggtagca   1050
                     M  G  W  S  C  I  I  L  F  L  V  A
                     Leader start

1051  acagctacaggtgtccactccgagctcacccagtctccatcctccctgtc   1100
       T  A  T  G  V  H  S  E  L  T  Q  S  P  -   (SEQ ID NO: 18)
                  Processed N-term.

1101  tgcatctgtaggagacagagtcaccatcacttgccgggcaactcagagtg   1150

1151  ttagtaactttttaaattggtatcagcagaagccaggggaagcccctacg   1200

1201  ctcctgatctatgatgcatccacttcgcaaagtggggtcccatcaaggtt   1250

1251  cagtggcagtggatctgggatggatttcagtctcaccatcagcagtctgc   1300

1301  agcctgaagatcttgcaatgtattactgtcaagcgagtatcaataccccg   1350

1351  cttttcggcggagggaccagaatagatatgagacgaactgtggctgcacc   1400

1401  atctgtcttcatcttcccgccatctgatgagcagttgaaatctggaactg   1450

1451  cctctgttgtgtgcctgctgaataacttctatcccagagaggccaaagta   1500

1501  cagtggaaggtggataacgccctccaatcgggtaactcccaggagagtgt   1550
```

14/31

1551 cacagagcaggacagcaaggacagcacctacagcctcagcagcaccctga 1600

1601 cgctgagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtc 1650

1651 acccatcagggcctgagcttgcccgtcacaaagagcttcaacaggggaga 1700
L P V T K S F N R G E

Xba I

1701 gtgttagtgagatgatcctctagagtcatctacgtatgatcagcctcgac 1750
C * end of light chain (SEQ ID NO: 19)

1751 tgtgccttctagttgccagccatctgttgtttgcccctcccccgtgcctt 1800

1801 ccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgag 1850

1851 gaaattgcatcgcattgtctgagtaggtgtcattctattctggggggtgg 1900

1901 ggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatg 1950

1951 ctggggatgcggtgggctctatggaaccagctggggctcgacagctcgag 2000

2001 ctagcttgcttctcaatttcttatttgcataatgagaaaaaaaggaaaa 2050

2051 ttaattttaacaccaattcagtagttgattgagcaaatgcgttgccaaaa 2100

2101 aggatgctttagagacagtgttctctgcacagataaggacaaacattatt 2150

2151 cagagggagtacccagagctgagactcctaagccagtgagtggcacagca 2200

2201 ttctagggagaaatatgcttgtcatcaccgaagcctgattccgtagagcc 2250

2251 acaccttggtaagggccaatctgctcacacaggatagagagggcaggagc 2300

2301 cagggcagggcatataaggtgaggtaggatcagttgctcctcacatttgc 2350

2351 ttctgacatagttgtgttgggagcttggatcgatccaccatggttgaaca 2400

2401 agatggattgcacgcaggttctccggccgcttgggtggagaggctattcg 2450

2451 gctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttc 2500

2501 cggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtc 2550

2551 cggtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctgg 2600

2601 ccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcg 2650

2651 ggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtc 2700

2701 atctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgc 2750

2751 ggcggctgcatacgcttgatccggctacctgcccattcgaccaccaagcg 2800

2801 aaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcga 2850

2851 tcaggatgatctggacgaagagcatcaggggctcgcgccagccgaactgt 2900

2901 tcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgtgacc 2950

2951 catggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttc 3000

3001 tggattcatcgactgtggccggctgggtgtggcggaccgctatcaggaca 3050

3051 tagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggct 3100

3101 gaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcat 3150

3151 cgccttctatcgccttcttgacgagttcttctgagcgggactctggggtt 3200

3201 cgaaatgaccgaccaagcgacgcccaacctgccatcacgagatttcgatt 3250

3251 ccaccgccgccttctatgaaaggttgggcttcggaatcgttttccgggac 3300

3301 gccggctggatgatcctccagcgcggggatctcatgctggagttcttcgc 3350

3351 ccaccccaacttgtttattgcagcttataatggttacaaataaagcaata 3400

3401 gcatcacaaatttcacaaataaagcatttttttcactgcattctagttgt 3450

3451 ggtttgtccaaactcatcaatgtatcttatcatgtctggatcgcggccgc 3500

3501 gatcccgtcgagagcttggcgtaatcatggtcatagctgtttcctgtgtg 3550

3551 aaattgttatccgctcacaattccacacaacatacgagccggaagcataa 3600

3601 agtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcg 3650

3651 ttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgca 3700

3701 ttaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgct 3750

3751 cttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcgg 3800

3801 cgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatc 3850

3851 aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca 3900

3901 ggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccc 3950

3951 cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaaccc 4000

4001 gacaggactataaagataccaggcgtttccccctggaagctccctcgtgc 4050

4051 gctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctc 4100

4101 ccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcag 4150

4151 ttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccg 4200

4201 ttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaac 4250

4251 ccggtaagacacgacttatcgccactggcagcagccactggtaacaggat 4300

4301 tagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggc 4350

4351 ctaactacggctacactagaaggacagtatttggtatctgcgctctgctg 4400

4401 aagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca 4450

4451 aaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgc 4500

4501 gcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtct 4550

4551 gacgctcagtggaacgaaaactcacgttaagggattttggtcatgagatt 4600

4601 atcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtttta 4650

4651 aatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgc 4700

4701 ttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccat 4750

4751 agttgcctgactccccgtcgtgtagataactacgatacgggagggcttac 4800

4801 catctggccccagtgctgcaatgataccgcgagacccacgctcaccggct 4850

4851 ccagatttatcagcaataaaccagccagccggaagggccgagcgcagaag 4900

4901 tggtcctgcaactttatccgcctccatccagtctattaattgttgccggg 4950

4951 aagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgcc 5000

5001 attgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcatt 5050

5051 cagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgt 5100

5101 gcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaag 5150

5151 ttggccgcagtgttatcactcatggttatggcagcactgcataattctct 5200

5201 tactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaa 5250

5251 ccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccg 5300

5301 gcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgct 5350

5351 catcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgc 5400

5401 tgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttca 5450

5451 gcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggca 5500

5501 aaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactca 5550

5551 tactcttcctttttcaatattattgaagcatttatcagggttattgtctc 5600

5601 atgagcggatacatatttgaatgtatttagaaaaataaacaaataggggt 5650

5651 tccgcgcacatttccccgaaaagtgccacct 5681 (SEQ ID NO: 17)

Figure 4C Heavy chain coding sequence in the plasmid Hu19BHepcd

```
                                                 Eco R1
                                                 gaattcggtacc    1000

1001  atggagtttgggctgagctgggttttcctcgtggctcttttaagaggtgt   1050
        M  E  F  G  L  S  W  V  F  L  V  A  L  L  R  G  V
       Leader start

1051  ccagtgtcaggtgcagctggtggagtctgggggaggcctggtcaggcctg   1100
        Q  C  Q  V  Q  L  V  - (SEQ ID NO: 21)
             Processed N-term

1101  gcgggtccctaagactctcgtgtgcagcctctggattcaccctcagtggc   1150

1151  tataccatgcactgggtccgccaggctccagggaaggggctggagtgggt   1200

1201  ctcatccattactggaggtagcaacttcatatactattcagactcagtga   1250

1251  agggccgattcaccatctccagagacaatgccaagaactcactttatctg   1300

1301  caaatgaacagcctgacagccgaggacacggctgtctattattgtgcgac   1350

1351  cgcccctatagcaccgccctactttgaccactggggccagggaaccctgg   1400

1401  tcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggca   1450

1451  ccctcctccaagagcacctctgggggcacagcggccctgggctgcctggt   1500

1501  caaggactacttccccgaaccggtgacggtgtcgtggaactcaggcgccc   1550

1551  tgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactc   1600
```

20/31

1601 tactccctcagcagcgtggtgaccgtgccctccagcagcttgggcaccca 1650

1651 gacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggaca 1700

1701 agaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgc 1750

1751 ccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaa 1800

1801 acccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtgg 1850

1851 tggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtg 1900

1901 gacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagta 1950

1951 caacagcacgtaccgggtggtcagcgtcctcaccgtcctgcaccaggact 2000

2001 ggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctccca 2050

2051 gcccccatcgagaaaaccatctccaaagccaaagggcagccccgagaacc 2100

2101 acaggtgtacaccctgcccccatcccgggatgagctgaccaagaaccagg 2150

2151 tcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtg 2200

2201 gagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcc 2250

2251 cgtgctggactccgacggctccttcttcctctacagcaagctcaccgtgg 2300

2301 acaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcat 2350

2351 gaggctctgcacaaccactacacgcagaagagcctctccctgtctccggg 2400

S P G

2401 taaatgatagatatc - (SEQ ID NO:20)

K * end of heavy chain

Figure 4D Light chain coding sequence in the plasmid Hu1961.cpen

Eco RI

<u>gaattc</u>**catgga** 1000

1001 **c<u>atg</u>agggtccccgctcagctcctagggctcctgctgctctggctccgag** 1050

M R V P A Q L L G L L L L W L R

Leader start

1051 **gtgccagatgtgacatccagatga**cccagtctccatcctccctgtctgca 1100

G A R C <u>D I Q</u> M T - (SEQ ID NO: 23)

Processed N-term

1101 tctgtaggagacagagtcaccatcacttgccgggcaactcagagtgttag 1150

1151 taactttttaaattggtatcagcagaagccagggaaagcccctaagctcc 1200

1201 tgatctatgatgcatccaatttgcaaagtggggtcccatcaaggttcagt 1250

1251 ggcagtggatctgggatggatttcagtctcaccatcagcagtctgcagcc 1300

1301 tgaagatcttgcaatgtattactgtcaagcgagtatcaataccccgcttt 1350

1351 tcggcggagggaccagaatagatatgagacgaactgtggctgcaccatct 1400

1401 gtcttcatcttcccgccatctgatgagcagttgaaatctggaactgcctc 1450

1451 tgttgtgtgcctgctgaataacttctatcccagagaggccaaagtacagt 1500

1501 ggaaggtggataacgccctccaatcgggtaactcccaggagagtgtcaca 1550

1551 gagcaggacagcaaggacagcacctacagcctcagcagcaccctgacgct 1600

1601 gagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtcaccc 1650

```
1651  atcagggcctgagcttgcccgtcacaaagagcttcaacaggggagagtgt   1700
                        L  P  V  T  K  S  F  N  R  G  E  C  (SEQ NO: 24)

                Xba I
1701  tagtgagatgatcctctaga          (SEQ ID NO: 22)
       *  end of light chain
```

Figure 4E Heavy chain coding sequence in the plasmid Hu19CHepcd

```
                                                 Eco RI
                                                 gaattcggtacc         1000

1001  atggagtttgggctgagctgggttttcctcgtggctcttttaagaggtgt    1050
       M   E   F   G   L   S   W   V   F   L   V   A   L   L   R   G   V

1051  ccagtgtcaggtgcagctggtggagtctgggggaggcctggtcaggcctg    1100
         Q   C   Q   V   Q   L   V  - (SEQ ID NO: 21)
                 Processed N-term

1101  gcgggtccctaagactctcgtgtgcagcctctggaaccaccctcagtggc    1150

1151  tataccatgcactgggtccgccaggctccagggaaggggctggagtgggt    1200

1201  ctcatccattactggcggtagcaacttcataaactacgcagactcagtga    1250
                               S   N   F   I   N   Y   A   (SEQ ID NO: 26)

1251  agggccgattcaccatctccagagacaacgccaagaactcactttatctg    1300

1301  caaatgaacagcctgagagccgaggacacggctgtctattattgtgcgac    1350

1351  cgcccctatagcaccgccctactttgaccactggggccagggaaccctgg    1400

1401  tcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggca    1450

1451  ccctcctccaagagcacctctgggggcacagcggccctgggctgcctggt    1500

1501  caaggactacttccccgaaccggtgacggtgtcgtggaactcaggcgccc    1550

1551  tgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactc    1600

1601  tactccctcagcagcgtggtgaccgtgccctccagcagcttgggcaccca    1650
```

1651 gacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggaca 1700

1701 agaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgc 1750

1751 ccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaa 1800

1801 acccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtgg 1850

1851 tggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtg 1900

1901 gacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagta 1950

1951 caacagcacgtaccgggtggtcagcgtcctcaccgtcctgcaccaggact 2000

2001 ggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctccca 2050

2051 gcccccatcgagaaaaccatctccaaagccaaagggcagccccgagaacc 2100

2101 acaggtgtacaccctgcccccatcccgggatgagctgaccaagaaccagg 2150

2151 tcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtg 2200

2201 gagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcc 2250

2251 cgtgctggactccgacggctccttcttcctctacagcaagctcaccgtgg 2300

2301 acaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcat 2350

2351 gaggctctgcacaaccactacacgcagaagagcctctccctgtctccggg 2400

S P G

2401 taaatgatagatatc - (SEQ ID NO: 25)

K * end of heavy chain

Figure 4F Heavy chain coding sequence in the plasmid Hu19DHeped

```
                                                   Eco RI
                                                   gaattcggtacc      1000

1001  atggagtttgggctgagctgggttttcctcgtggctcttttaagaggtgt       1050
       M   E   F   G   L   S   W   V   F   L   V   A   L   L   R   G   V

1051  ccagtgtcaggtgcagctggtggagtctgggggaggcttggtccagcctg       1100
         Q   C   Q   V   Q   L   V   - (SEQ ID NO: 21)
                Processed N-term

1101  gggggtccctgagactctcctgtgcagcctctggaaccaccctcagtggc       1150

1151  tataccatgcactgggtccgccaggctccagggaaggggctggagtgggt       1200

1201  ctcatccattactggaggtagcaacttcatacaatactcagactcagtga       1250
                              S   N   F   I   Q   Y   S - (SEQ ID NO: 28)

1251  agggccgattcaccatctccagagacaacgccaagaactcactttatctg       1300

1301  caaatgaacagcctgacagccgaggacacggctgtctattattgtgcgag       1350

1351  agcccctatagcacggccctactttgaccactggggccagggaaccctgg       1400

1401  tcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggca       1450

1451  ccctcctccaagagcacctctgggggcacagcggccctgggctgcctggt       1500

1501  caaggactacttccccgaaccggtgacggtgtcgtggaactcaggcgccc       1550

1551  tgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactc       1600

1601  tactccctcagcagcgtggtgaccgtgccctccagcagcttgggcaccca       1650
```

1651 gacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggaca 1700

1701 agaaagttgagcccaaatcttgtgacaaaactcacacatgcccaccgtgc 1750

1751 ccagcacctgaactcctggggggaccgtcagtcttcctcttccccccaaa 1800

1801 acccaaggacaccctcatgatctcccggacccctgaggtcacatgcgtgg 1850

1851 tggtggacgtgagccacgaagaccctgaggtcaagttcaactggtacgtg 1900

1901 gacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagta 1950

1951 caacagcacgtaccgggtggtcagcgtcctcaccgtcctgcaccaggact 2000

2001 ggctgaatggcaaggagtacaagtgcaaggtctccaacaaagccctccca 2050

2051 gcccccatcgagaaaaccatctccaaagccaaagggcagccccgagaacc 2100

2101 acaggtgtacaccctgcccccatcccgggatgagctgaccaagaaccagg 2150

2151 tcagcctgacctgcctggtcaaaggcttctatcccagcgacatcgccgtg 2200

2201 gagtgggagagcaatgggcagccggagaacaactacaagaccacgcctcc 2250

2251 cgtgctggactccgacggctccttcttcctctacagcaagctcaccgtgg 2300

2301 acaagagcaggtggcagcaggggaacgtcttctcatgctccgtgatgcat 2350

2351 gaggctctgcacaaccactacacgcagaagagcctctccctgtctccggg 2400

S P G

2401 taaatgatagatatc (SEQ ID NO: 27)

K * end of heavy chain

Figure 4G Light chain coding sequence in the plasmid H19CLopen

Eco RI

gaattccatgga 1000

1001 catgagggtccccgctcagctcctggggctcctgctgctctggctccgag 1050

M R V P A Q L L G L L L L W L R

Leader start

1051 gtgccagatgtgacatccagatgacccagtctccatcctccctgtctgca 1100

G A R C D I Q M T - (SEQ ID NO: 23)

Processed N-term

1101 tctgtaggagacagagtcaccatcacttgccgggcaactcagagtgttag 1150

1151 taactttttaaattggtatcagcagaagccaggggaagcccctaagctcc 1200

1201 tgatctatgatgcatccacttcgcaaagtggggtcccatcaaggttcagt 1250

1251 ggcagtggatctgggatggatttcactctcaccatcagcagtctgcagcc 1300

1301 tgaagatcttgcaatgtattactgtcaacagagtatcaataccccgcttt 1350

1351 tcggcggagggaccagaatagatatgagacgaactgtggctgcaccatct 1400

1401 gtcttcatcttcccgccatctgatgagcagttgaaatctggaactgcctc 1450

1451 tgttgtgtgcctgctgaataacttctatcccagagaggccaaagtacagt 1500

1501 ggaaggtggataacgccctccaatcgggtaactcccaggagagtgtcaca 1550

1551 gagcaggacagcaaggacagcacctacagcctcagcagcaccctgacgct 1600

```
1601  gagcaaagcagactacgagaaacacaaagtctacgcctgcgaagtcaccc   1650

1651  atcagggcctgagctcgcccgtcacaaagagcttcaacaggggagagtgt   1700
                      S  P  V  T  K  S  P  T  R  G  Q  C  (SEQ NO: 30)

                  Xba I
1701  tagtgagatgatcctctagatctacgtatgatcagcctcgactgtgcctt -(SEQ NO: 29)
        *  end of light chain
```

Fig. 5. COOMASSIE STAINED SDS-PAGE GEL ANALYSIS OF 10UG HU19B AND HU19C RESPECTIVELY UNDER REDUCING CONDITIONS

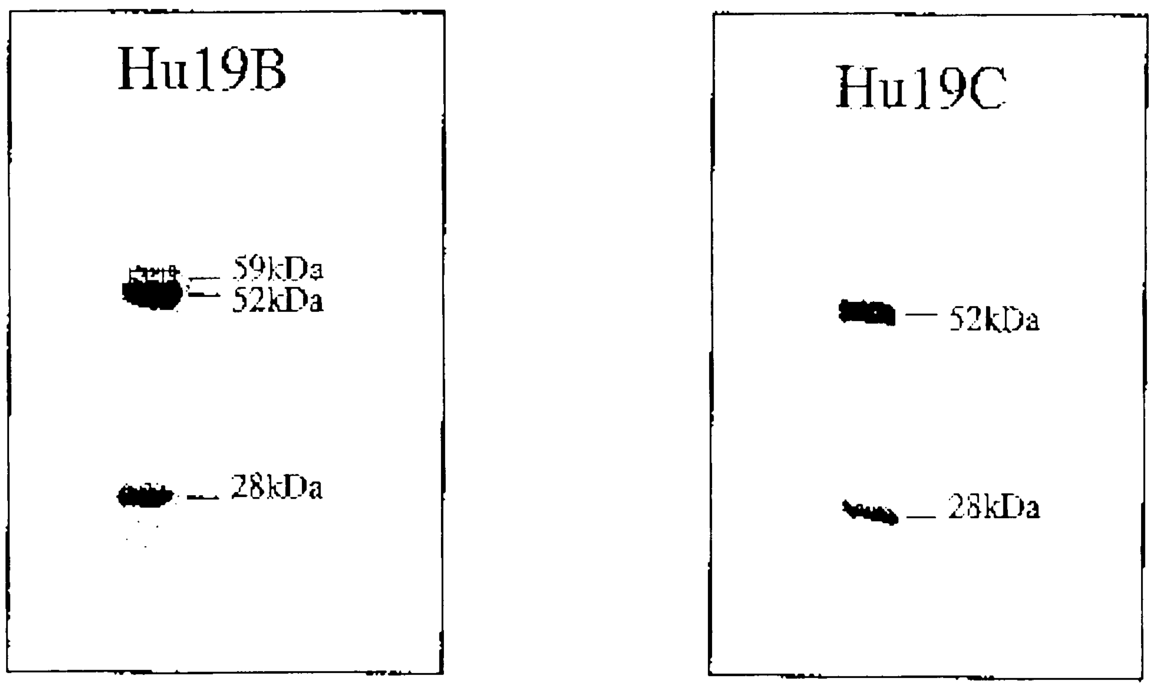

Fig. 6. SEPARATION OF HU19B GLYCOVARIANTS BY ANION EXCHANGE CHROMATOGRAPHY

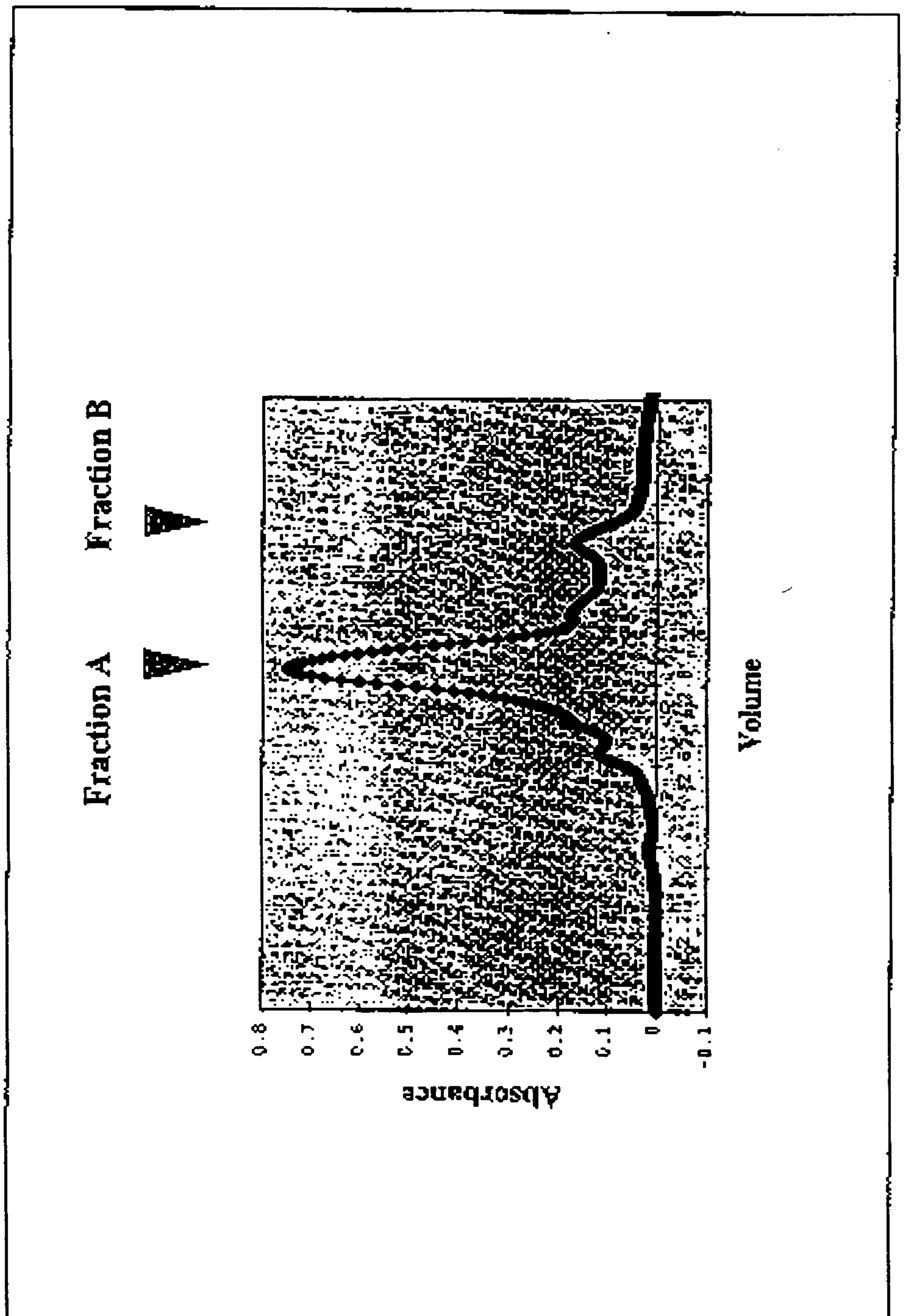

Fig.7. SDS-PAGE of Glycosylation Variant of Hu19B Fab